AD NUMBER

AD866166

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; FEB 1970. Other requests shall be referred to Rome Air Development Center, Attn: EMNRR, Griffiss AFB, NY 13440.

AUTHORITY

RADC USAF ltr, 15 Dec 1971

BAYESIAN RELIABILITY DEMONSTRATION: PHASE I - DATA FOR THE Á PRIORI DISTRIBUTION

R. E. Schafer
John Collins
M. L. Luden
et al

Hughes Aircraft Company

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of RADC (EMNRR), GAFB, N. Y. 13440.

FOREWORD

This final report was prepared by John Collins, M. L. Luden, Dr. Nozer Singpurwalla, Tom Sheffield and Ray E. Schafer of Hughes Aircraft Company, Ground Systems Group, Fullerton, California under Contract Number F30602-69-C-0042, Project 5519, Task 551902. The RADC Project Engineer was Anthony Feduccia (EMNRR).

This technical report has been reviewed and is approved.

Approved: Anthony J. Feduccia
ANTHONY J. FEDUCCIA
Reliability Engineering Section
Reliability Branch

Approved: Joseph J. Naresky
JOSEPH J. NARESKY
Chief, Reliability & Compatibility Division

FOR THE COMMANDER: Irving J. Gabelman
IRVING J. GABELMAN
Chief, Plans Office

ABSTRACT

This final report is a result of a study performed for RADC under Contract Number F30602-69C-0042. The purpose of the study was to fit one or more prior distributions to θ = MTBF = Mean Time Between Failure. In particular, the objectives were three:

i) Establish criteria for data that would be suitable for fitting prior distributions to θ = MTBF.

ii) Develop methods of fitting and fit one or more prior distributions.

iii) Perform robustness analysis of fitted prior distributions.

It was discovered that if the number of identical equipments and number of failures observed per equipment are relatively small, special methods of fitting are required. For the data used in this study, the inverted gamma distribution turned out to be a good prior fit.

EVALUATION

A thorough review of the available literature on Bayesian reliability (as listed in Section 11 of this report) will reveal that very little, if any, effort has been directed towards developing methods for fitting prior distributions to empirical data. Usually, priors are determined through engineering judgment, personal experience, and wide-scale assumptions. Such means of developing priors are unsatisfactory if Bayes' methods are to be used in equipment reliability demonstration. The objectives of this Phase I study were to (a) establish data criteria for use in fitting priors, (b) develop methods of fitting and actually fit one or more priors, and (c) perform robustness analysis of fitted priors. The results of Phase I show, for the first time, how priors can be fitted to empirical data, the amount and type of data required, and the effect on the posterior distribution of varying parameters of the prior distribution.

Each of the objectives of Phase I was successfully completed. Section 5 shows, for three forms of operational data, the minimum values of "n" (number of equipments) and "k" (number of failures per equipment) required for fitting a valid prior distribution to MTBF. The results of this section provide a firm foundation on which future data collection programs for fitting priors can be based. Section 4 presents methods of fitting prior distributions and includes seven inverted gamma priors fitted to empirical data collected on seven different types of equipment. These results confirm the practicality of the Bayes' method in reliability demonstration, and provide justification for the use of the inverted gamma as the prior distribution on equipment MTBF. Section 7 contains the results of the robustness analysis performed to investigate the effects of errors in estimating the scale and shape parameters of the inverted gamma prior on the posterior inverted gamma distribution. In addition, as previously stated, Section 11 includes a comprehensive, 78-entry bibliography of sources on Bayes' reliability.

The methods developed in this study were based on "equipment" level data, thus it was assumed that the conditional distribution of time-to-failure was assumed to be exponential. However, it should be noted that the methods given here are generally applicable whatever the form of the conditional distribution.

Phase I results will serve as important inputs to the Phase II study which is scheduled to start in February 1970. The methods developed and data criteria established in Phase I will be used to fit additional priors on other equipment types in Phase II. In addition, this next phase will include the investigation of methods of combining priors from similar, but not identical, equipment. Phase II objectives also include the establishment of plans for developing and implementing Bayesian reliability demonstration tests, as a prelude to the projected FY-71 Phase III study.

Anthony J. Feduccia

ANTHONY J. FEDUCCIA
Reliability Engineering Section
Reliability Branch

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE		PAGE

LIST OF TABLES

0.0 SUMMARY

The knowledge of the prior distribution is central to the efficient use of Bayes methods in reliability estimation and demonstration. This study presents methods and examples of fitting prior distributions when the data available is:

i) n sample MTBF's on n identical equipments.

ii) n observations on the number of failures occurring in a fixed time T for n identical equipments.

This data must be used since samples from the prior distribution itself, i.e., the true MTBF, are unobtainable. Two cases were considered: the family of the prior distribution being specified and unspecified. In case the analyst is unwilling to specify the family the methods are not wholly satisfactory; the most suitable being a method due to Krutchkoff and Rutherford which requires the prior distribution belong to the Pearson class. When the family is specified (e.g., inverted gamma) methods are presented which use the known marginal distribution to fit the prior distribution. Some difficulty is experienced when the sample sizes (K_i) for the n sample MTBF's are not all identical and small. In this case a mixed model for the marginal distribution is used. If the $\{K_i\}$ are all large (roughly > 30) even though not identical the marginal distribution approaches the prior distribution rapidly. If n and $\{K_i\}$ are both large the data is fitted directly to the prior distribution.

The results are given in the following table:

General Methods of Fitting Prior Distributions

Family Unspecified (all values of n and $\{K_i\}$).	Estimate first four moments of prior distribution from data and use Pearson class.
Family Specified $\{K_i\}$ large, n large enough to use χ^2 test.	Treat data as though it came from prior distribution and use classical methods.
$\{K_i\}$ small, n large enough to use mixed model χ^2 test.	Fit data to known marginal using classical methods.

In the text, specific results are given as to the meaning of "large".

n = number of identical equipments.

K_i = number of failures observed on the i^{th} equipment.

1.0 INTRODUCTION

1.1 OBJECTIVE OF THE STUDY

There are a number of measures of reliability. The most important are

i) Mean Time Between Failure (MTBF).

ii) Failure Rate.

iii) Probability of Survival for a fixed time T.

iv) Time, say x_p, for which the probability of survival is p. That is, x_p is the $(1-p)^{th}$ quantile of the time to failure distribution.

It is customary in equipment/system development to place certain specifications on one or more of the above measures. Demonstration tests (in statistical language hypothesis tests) are a fundamental tool in verifying that these specifications have been met. Unfortunately, the demonstration tests are conducted in an environment of

- High reliability requirements.
- Limited funds.
- Short time available.

Generally, the precision of the demonstration test increases as the number of failures observed increases. But high reliability requirements mean long times to observe failures. On the other hand, low producer and consumer risks are desired and this also means long test times. This creates a cost/time problem which is apparently unsolvable by classical methods. In fact, Bayes methods apparently hold the most promise in solving the cost/time problem of demonstration tests.

In many sampling situations, Bayes methods are not applicable because the unknown parameter (here, MTBF, failure rate, etc.) cannot be considered a random variable, which is a "must" to use Bayes methods. However, it is clear that in reliability the parameter can often, if not always, be considered a random variable. Consider a computer manufactured by a particular Company to a particular design; each successive computer differing in serial number and parts. Because of these part (not part type) differences and other differences, each computer will possess a different true (but unknown) MTBF = θ and hence, MTBF = θ may be considered a random variable. The essence of Bayes methods is that a probability distribution is assumed to exist on the parameter (here, measure of reliability) in question. This probability distribution is called the prior distribution.*

The measure of reliability considered throughout this study is i) above, namely, the very important MTBF = θ. Not only is MTBF most commonly used as a measure of reliability, it has the advantage that it is a parameter in the time to failure distribution, which results in variables type data rather than iii) and iv) above, which result in attributes data. Thus, MTBF permits a parametric rather than a nonparametric approach.

In summary then, to use Bayes methods a prior distribution is needed. The basic objective of this study is to fit one or more prior distributions to MTBF = θ for ground electronic equipments. In doing this, there are many related questions so that the study objectives are more particularly stated as:

i) Establish criteria for suitable data for fitting prior distributions to θ,

ii) Fit one or more prior distributions.

iii) Perform robustness analysis of fitted prior distributions.

All these objectives were accomplished and are discussed in the appropriate sections.

The next section gives more detail and insight into the use of Bayes methods in reliability.

1.2 INTRODUCTION TO BAYES METHODS IN RELIABILITY

A classical lower one sided confidence interval for θ (MTBF) consists of a statement that

$P(\theta \geq \theta_*) = 1-\gamma$, γ small and is prepared after the data is available. The probability can only be interpreted as a confidence. A classical (e.g., MIL-STD-781B) reliability demonstration test consists of preselecting (i.e., before the data is gathered)

$\theta_* = \theta_1$ (minimum acceptable MTBF) and β = consumer's risk. In order to determine a unique test it is also required to "add" a θ_0 (specified MTBF) and α = producer's risk. It is interesting and informative to dwell on the "names" of θ_1 and θ_0, i.e., minimum acceptable and specified respectively. One might think, in view of the "name" of θ_1 that θ_0 would be called maximum acceptable or simply acceptable θ. If this were so, things would be very misleading. That is, θ_0, θ_1 are well named now. The name for θ_1 depicts what it really is: a prespecified lower confidence limit. The name for θ_0 depicts what it really is: an "add on" to obtain a unique test. Briefly, it is consumers who demand tests. They specify (θ_1, β). Once having done this the producer selects the (unique) test suitable to him.

*In the more mathematical literature, this prior distribution is sometimes called the Compounding or Mixing distribution.

Using Bayes methods it is possible to prepare a probability (as against confidence) statement to the effect

$$P(\theta \geq \theta_* \mid \text{observed data}) = 1 - \gamma \tag{1.2.1}$$

A unique Bayes demonstration test can be achieved with various "add ons" but that is not the subject of this report. The important point is the conditioning random variable "observed data" in (1.2.1) above. It is immaterial what it is as long as it is a sufficient statistic for θ. Sometimes the vector of failure times $(x_1, \ldots, x_K)$ may be used and sometimes the observed MTBF $\hat{\theta}$ may be used. In any event to prepare (1.2.1), i.e., a Bayes confidence interval,

$$P(\theta \geq \theta_* \mid \text{observed data} = \hat{\theta}) = \int_{\theta_*}^{\infty} g(\theta|\hat{\theta})\, d\theta,$$

the posterior distribution of θ is required. Here we have illustrated it and will continue to illustrate with the conditioning random variable $\hat{\theta}$.

Clearly,

$$g(\theta|\hat{\theta}) = \frac{f(\hat{\theta}|\theta)g(\theta)}{f(\hat{\theta})}\ ; \ f(\hat{\theta}) \neq 0. \tag{1.2.2}$$

Thus, to obtain a Bayes confidence interval one must have

i) The prior distribution $g(\theta)$.

ii) The sampling distribution of the conditioning statistic given θ, i.e., $f(\hat{\theta}|\theta)$.

iii) The marginal distribution of the statistic $f(\hat{\theta})$.

Each of these distributions, particularly i) and iii) above, play a vital role in this report. The object is to fit $g(\theta)$. The data gathered for this study is at the equipment level and it is assumed throughout that the conditional distribution of times to failure given θ, i.e., $f(x|\theta)$, is exponential so that $f(\hat{\theta}|\theta)$ is gamma. It is then relatively simple to find $f(\hat{\theta})$ once the prior distribution is specified since

$$f(\hat{\theta}) = \int_0^{\infty} f(\hat{\theta}|\theta)g(\theta)\, d\theta \tag{1.2.3}$$

In some cases the observed random variable is the number of failures (say x) occurring in a fixed time period T so that the $f(x|\theta)$ is Poisson and the marginal distribution of x is

$$f(x) = \int_0^\infty f(x|\theta)g(\theta)\, d\theta \qquad (1.2.4)$$

Thus, of the four (4) distributions of (1.2.2), only two (2) are unknown in the sense they must be estimated since the marginal distribution in the denominator is obtainable through $g(\theta)$ and $f(\hat{\theta}|\theta)$.

1.3 THE CENTRAL PROBLEM

The central objective of this study is to fit a prior distribution to some reliability data; the central problem is that for a given piece of equipment the true θ is unknown and remains unknown unless an infinite number of failures are observed. Thus, in fitting $g(\theta)$, one does not even have random samples from $g(\theta)$. To obtain a random sample from $g(\theta)$ would be to know the MTBF of an equipment exactly which is impossible. There are then two sample sizes of concern. First, the number of "identical" equipments sampled which we call, hereafter, n. Second, there is the number of failure times available on each equipment which, for the i^{th} equipment, we call K_i and for all n equipments we denote the set of K_i as $\{K_i\}_n$. The situation is depicted below in Figure 1.3. The squares represent identical equipments. In this case x represents lifetime.

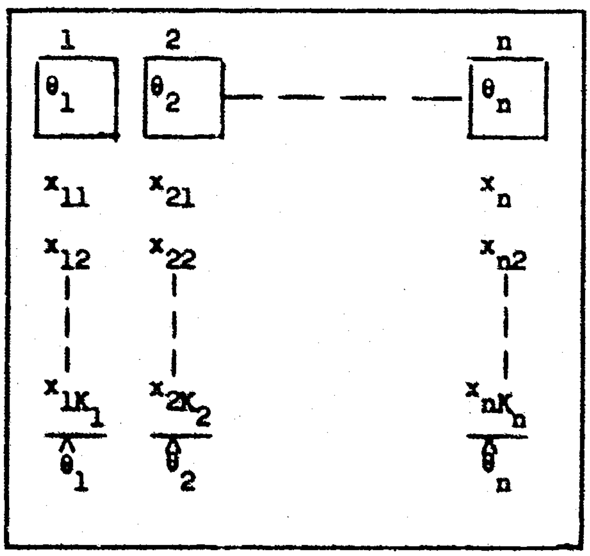


Figure 1.3 Variables Involved in Fitting a Prior Distribution

At first thought, one might be tempted to plot the $\hat{\theta}$'s (somehow weighted for unequal K's) in a histogram and by usual methods fit some prior distribution. To see that this is not correct, consider first the case of identical K's. If the $\hat{\theta}$'s were plotted in a histogram and n taken very large the histogram would look very much like $f_K(\hat{\theta})$. We are essentially "summing" over θ in the joint density $h_K(\hat{\theta},\theta)$. That is, each sample involves a pair $(\hat{\theta}_i,\theta_i)$ $i=1, \ldots, n$ but the θ_i component is unobservable. Thus, each $\hat{\theta}$ may be considered a random sample from $f_K(\hat{\theta})$. In the case of unequal K's (but not all necessarily different) things get messy quickly but a mixed population model appears descriptive, i.e., each $\hat{\theta}$ is regarded as a random sample from

$$f_{\{K_i\}_n}(\hat{\theta}) = p_1 f_{K_1}(\hat{\theta}) + \ldots + p_{K_t}\, f_{K_t}(\hat{\theta}) \tag{1.3.1}$$

Where t is the number of distinct K_i's and $p_i = \dfrac{\text{number of identical } K_i\text{'s}}{n}$ also $\Sigma p_i = 1$. The numerator in p_i is taken after renumbering to obtain different K_i's. This model will be discussed in more detail in Section 4.0.

The important point of this whole discussion is:

> Whatever methods are developed for fitting the prior distribution, the data to be used will be data which is taken from the marginal distribution not from the prior distribution.

The next section gives definitions which will be used throughout this report.

1.4 DEFINITIONS AND ASSUMPTIONS

Symbol	Definition
n	The number of identical equipments used to fit a prior distribution.
K_i	The number of failures available on the i^{th} equipment. $i=1, \ldots, n$.
$\hat{\theta}_i$	$\dfrac{\text{Total test time for } i^{th} \text{ equipment}}{K_i}$
θ	The unknown MTBF of an equipment.
X	The random variable denoting either time to failure or number of failures in a fixed time. The context will make clear which is being discussed.
x	A particular value of X.

Symbol	Definition
$\chi^2_{q,u}$	The q^{th} quantile of the Chi-Square distribution with u degrees of freedom.
$g(\theta)$	The prior distribution on the random variable θ.
$g(\theta\|\hat{\theta})$	The posterior (conditional) distribution of θ having observed $\hat{\theta}$.
$f(\hat{\theta}\|\theta)$	The conditional distribution of $\hat{\theta}$ for fixed θ.
$f_K(\hat{\theta})$	The marginal (or unconditional) distribution of $\hat{\theta}$ where $\hat{\theta}$ is based on K failures.
$f(x)$	The marginal (or unconditional) distribution of time to failure or number of failures per fixed time T. The context will make it clear which is being discussed.
$f(x\|\theta)$	The conditional distribution of time to failure (or failures per T) for fixed θ.

One of the important assumptions of this report is that the conditional distribution of time to failure for fixed θ is exponential:

$$f(x|\theta) = 1/\theta \; e^{-x/\theta} \qquad x,\theta > 0. \tag{1.4.1}$$

Hence, the number of failures occurring in fixed time T is Poisson:

$$f(x|\theta) = \frac{e^{-T/\theta}(T/\theta)^x}{x!} \qquad T,\theta > 0 \quad x = 0,1,\ \ldots \tag{1.4.2}$$

The ultimate aim of fitting prior distributions in reliability is Bayes demonstration tests. Since demonstration tests are most generally applied at equipment/system levels rather than part levels the study was restricted to this level of data. For this reason, the exponential assumption was felt to be valid since it has been experienced repeatedly and certain limit theorems (e.g., the Drenick theorem) indicate an exponential distribution is to be expected for complex equipment. However, the methods developed here are generally applicable whatever the form of the conditional distribution. What might happen though for other conditional distributions is that the arithmetic will rapidly become intractable.

SECTION 2.0 DATA COLLECTION

2.1 TYPES OF DATA

Failure data can occur in several forms. It is divided into two main classes: attributes and variables data. The attributes situation occurs when an equipment is operated and its survival or nonsurvival for time T (usually mission time) is observed. Due to the nature of this type of data, attributes data was considered unusable for this study, therefore, no data of this type was collected.

The variables data situation occurs when the actual failure times are available. These times occur by agreeing to stop testing either after a fixed number of failures have occurred or after a fixed time has elapsed. The former case is called a censored test and the latter is called a truncated test. Often the failure times themselves are not available but the pair ($\hat{\theta}$ = observed MTBF, K = number of failures) is available.

2.2 DATA COLLECTION PLAN

The search for potential data sources was initially limited to Hughes Aircraft and Government sources. A complete and comprehensive search for high quality reliability data on ground electronic equipment was conducted in all of Hughes Aircraft Divisions including the Quality, Reliability, Effectiveness, Field Service Organizations and the major Government data centers maintained by the U. S. Government Agencies. Much data was available but due to the high quality restriction: equipment level with a fairly large, the amount of data useful for the establishment of a priori distributions was greatly reduced. The data location effort in the later phase of the search included contacts in private industry. Due to the nature of the data (MTBF) requested and proprietary rights of the contractor we could obtain no data from industry.

2.3 DATA COLLECTED

2.3.1 HUGHES INTERNAL SOURCES

The primary source of high quality data for establishing a priori distributions was Hughes Aircraft Company. The Hughes Aerospace Division has performed AGREE tests on various Hughes built electronic equipments. The Hughes Ground Systems Group has collected data on its systems and equipment for many years. This data includes both laboratory and field performance data.

Table 2.3.1.1 lists the data located and acquired within Hughes Aircraft for utilization in this study.

TABLE 2.3.1.1 HUGHES AIRCRAFT COMPANY ACQUIRED DATA

Assigned Data Set Number	Equipment	n	ΣK_i	Type of Tests
9.	Iram Computer	28	44	Lab (AGREE)
10.	Multi-mode Storage Tube Indicator (Display Tube)	20	24	Lab (AGREE)
11.	Infrared Subsystem in Fire Control Systems (MG-13)	42	106	Lab (AGREE)
12.	H-3118 Computers	43	2779	Lab/Field
13.	Display Console	48	314	Lab/Field
14.	H-3182 Converters D-A	19	102	Lab/Field
15.	Magnetic Tape Units	14	50	Lab/Field
16.	Infrared Subsystem in Fire Control Systems (MA-1)	43	99	Lab (AGREE)
17.	Rapid Tune Subsystem	12	41	Lab (AGREE)

2.3.2 GOVERNMENT DATA SOURCES

The search for and acquisition of data from the major data centers maintained by the U. S. Government Agencies was done by direct contact. Because of the vast amount of data collected and available in the Government data centers, the efforts were directed to collecting only adequate and useful data. It was felt that a visit to the key data centers was a necessity to assure a thorough analysis of the type of data collected and the correct data retrieval methods.

The agencies visited were:

1) U.S. Army Maintenance Command Logistic Data Center (USAMCLDC), Lexington, Kentucky (TAERS data).

2) U. S. Navy, Washington, D.C. (3M data).

3) U. S. Navy, Norfolk, Virginia, Statistical Engineering Branch, NAVSECNORDIV.

4) U. S. Naval Fleet Missile System Analysis and Evaluation Group (FMSAEG), Corona, California.

5) U. S. Air Force Logistics Command, Tinker Air Force Base, Oklahoma City, Oklahoma (66-1 data).

There was no reliability data available on electronic equipment at the USAMC Logistic Data Center due to the lack of control of serial number assignment to end items and usage. Indicating devices being almost non-existent on electronic equipment resulted in no usage time; therefore, no MTBF data is available through TAERS.

An additional request, suggested by the USAMCLDC, was made to the U. S. Army Electronics Command in Fort Monmouth, New Jersey, Applications Engineering Branch which resulted in MTBF data, on radio sets but was unsatisfactory (n too small) for use in establishing a priori distributions.

The investigation of the 3M data collection system maintained by the Navy yielded no reliability data; i.e., MTBF's. The Statistical Engineering Branch in Norfolk, Virginia had some reliability data but not of the quality of data necessary for this study. A trip to the U. S. Naval Fleet Missile Systems, Analysis and Evaluation Group in Corona, California was made in search of MTBF data on surface missile systems. Again, the quality (n too small) of data was unsatisfactory for the prior distributions requirement.

A visit to the Reliability Branch at Tinker Air Force Base in Oklahoma City, Oklahoma resulted in acquisition of reliability data on ground communications radars. Table 2.3.2.1 is a list of equipments from which 66-1 form data was obtained. All of this data was based on a fi d time T of 4320 hours.

TABLE 2.3.2.1 TINKER AIR FORCE BASE ACQUIRED DATA

Assigned Data Set Number	Equipment	n	ΣK_i
1.	MTI Reflector	41	173
2.	HVPS (High Voltage Power Supply)	74	475
3.	Synchronizer	59	265
4.	Oscilloscope	51	159
5.	Video Amplifier	40	92
6.	Synchronous Power Supply	55	183
7.	Search Indicator	58	266
8.	Servo Amplifier	50	194

2.3.3 DATA SOURCES

1. Interceptor Improvement Program IP Reliability Testing, Final Report, May 1964, Hughes Aircraft Company.

2. MMST Production Reliability Testing, "First Reliability Sampling Test Report," July 1967, "Second Reliability Sampling Test Report," October 1967, "Fourth and Final Reliability Sampling Test Report," December 1967, Hughes Aircraft Company.

3. MA-1/AN/ASQ-25/M6-13 Interceptor Improvement Program, "Rapid Tune Reliability Testing," November-December 1965, Hughes Aircraft Company.

4. Iram Production Reliability Tests, Final Test Report, June 1968, Hughes Aircraft Company.

2.3.4 SUMMARY OF DATA AVAILABILITY

It is well-known in statistical analysis that data can arise in two ways: i) after the fact, ii) as a result of an experiment designed expressly to answer specific questions; for example, to fit a prior distribution. The results of this study make it clear what to do if data is to be gathered expressly to fit prior distributions. However, if time and money are not available for the second approach above, already existing data must be used. The results of the data search indicate there is not much suitable data available. The reasons are primarily two

1) Not much data is available on large numbers of identical equipments.

2) Most data, already in existence, involves different numbers of failures or different (fixed) test time on each equipment.

The two reasons above are somewhat different in character. The first, primarily, causes poor fits while the second makes it difficult to apply the methods of this report at all.

Since large amounts of data are not available through government sources, the data for prior fits must come from industry. The picture may appear unnecessarily bleak: the results of Sections 4.0 and 5.0 indicate data requirements which are not too stringent.

SECTION 3.0 FUNCTIONAL LEVEL - THE CONDITIONAL DISTRIBUTION

In using Bayesian methods the prior distribution receives a good deal of attention - and rightly so. However, in any analysis involving actual data (X) the conditional distribution of $(X|\theta)$ must also be known. This conditional p.d.f., sometimes called a sampling distribution, receives less attention because, usually, more is known about it than about the prior distribution. For example, if several machines are turning out a large number of bolts which can be classified only as good or bad then for fixed fraction defective p the sampling distribution or conditional distribution is hypergeometric or binomial depending on the finiteness of the outputs. Thus, often the physical process dictates the sampling distribution. In other cases, experience has dictated what to expect for a conditional distribution. That is the case here. We have assumed that the important Bayes testing applications will be made at the equipment (computer, radar, oscilloscope, radio, etc.) or system level. In this case, it has been demonstrated many times both by limit theorems and empirical studies that if the random variable is time to failure, a good descriptor of the conditional distribution is the exponential, i.e.,

$$f(x|\theta) = 1/\theta \; e^{-x/\theta} \qquad x, \theta > 0$$

$$= 0 \text{ elsewhere.}$$

Then the distribution of the number of failures in fixed time T is Poisson, i.e.,

$$f(x|\theta) = \frac{e^{-T/\theta} (T/\theta)^x}{x!} \qquad t, \theta > 0$$

$$x = 0, 1, \ldots$$

$$= 0 \text{ elsewhere.}$$

It should be understood that the methods given here are illustrated and fully developed for the above two conditional distributions but that in general the methods are applicable to any conditional distribution provided the identifiability criterion is satisfied (see Appendix, Section 9.4).

If it should turn out that the conditional distribution is Weibull, then the same general methods would apply with the exception that the prior distribution would be two (2) variate. The major change then would be that the arithmetic would become much more intricate.

SECTION 4.0 METHODS OF FITTING PRIOR DISTRIBUTIONS AND RESULTS

SUMMARY

Objective

The objective of this section is to develop methods of fitting prior distributions to θ = MTBF when data occurs in any one of three (3) forms.

Type 1 - The recorded data on a piece of equipment is number of failures occurring in a fixed time T, T identical for all n equipments.

Type 2a - The recorded data on a piece of equipment is the observed MTBF, $\hat{\theta}$, and the number of failures on each identical equipment is a fixed K.

Type 2b - The recorded data is the same as Type 2a but the number of failures on each identical equipment may vary.

Two cases were considered: the family of the prior distribution unspecified and specified. These are discussed in more detail in Section 4.2.1

Results

Family of prior distribution unspecified, all types of data.

For the family completely unspecified the methods found were unsatisfactory for two reasons, the first being the most important.

1) They give only an empirical distribution for the prior and do not identify a family.

2) They are extremely complex to apply.

If one is willing to assume the prior distribution belongs to the Pearson family, the Krutchkoff-Rutherford method is easy to apply but requires an enormous number of identical equipment for good results.

Family of prior distribution specified, all types of data.

When the family is specified, the steps in fitting the prior distribution are simple enough.

1) Derive the marginal distribution under the assumed prior family.

2) Estimate the parameters of the marginal (and hence, prior) distribution.

3) Perform a χ^2 test of goodness of fit to the expected marginal distribution.

4) Assign appropriate prior distribution.

In this report a prior inverted gamma family was assumed. The appropriate marginal distributions are derived in Section 9.2. The appropriate steps are discussed in detail in Sections 4.2.2 and 4.2.3. The results are given in Section 4.3.2. Briefly, it was found that for type 1 data, of which there were eight (8) sets, seven (7) were found to be good inverted gamma fits. There was no type 2a data and for type 2b data fits could not be obtained because K (the number of failures available on each equipment) varied too much compared to the available number of units n.

4.1 FITTING THE PRIOR DISTRIBUTION WHEN THE FAMILY IS UNSPECIFIED

The case when the analyst is unwilling to assume any knowledge of the prior distribution is extremely difficult and remains, at this writing, essentially an unsolved problem. An idea why can be obtained by looking at the classical χ^2 test. Even in this test the family of the fitted distribution must be specified. The methods we will discuss here all have one point in common:

They all use the data $(x_1, \ldots, x_n)$ or $(\hat{\theta}_1, \ldots, \hat{\theta}_n)$ to estimate the MOMENTS of the prior distribution and then apply certain uniqueness theorems between a sequence of moments and a distribution.

The extent to which a sequence of moments determines a unique probability distribution AND how one is to discover, having a sequence of moments, to which probability distribution it corresponds is a difficult problem (see Shohat, J. A. and Tamarkin, J. D. (1943), The Problem of Moments, Amer. Math. Society, New York). The methods to be discussed assume some of the difficulties away.

One method of interest is given by John E. Rolph (Bayesian Estimation of Mixing Distributions, Annals of Math. Statistics, Vol. 39, No. 4, August 1968). In this paper Dr. Rolph assumes a prior distribution (uniform) on the family of prior distributions. Moreover, he assumes

i) The parameter space θ is limited to $[0,1]$, i.e., $0 \leq \theta \leq 1$.

ii) The conditional distribution $f(x|\theta)$ must be a polynomial in θ and be a discrete distribution.

The first assumption is needed to use the theorem that a probability distribution on $[0,1]$ is uniquely defined by its moments. The first assumption also permits the prior distribution on the family of prior distributions to be assigned to the moment sequences. The second assumption permits the marginal distribution

$$f(x) = \int_0^1 f(x|\theta)\, g(\theta)\, d\theta$$

to be written

$$f(x) = \sum_i a_{xi} \mu_i$$

where μ_i is the i^{th} moment of the prior distribution. Thus, a sample

$(x_1, \ldots, x_n)$ from $f(x)$ can be used to estimate the moments of the prior distribution and hence (because of uniqueness) estimate the prior distribution. The estimate constructed in this manner is consistent.* The required computations are quite involved, surely requiring a computer program since a number of relatively high order determinants are involved. This method is not directly useful here since the restriction of θ to $[0,1]$ is untenable. However, it is mentioned since, in the exponential case, the reliability function

$$R(T) = e^{-T/\theta} \quad \underline{T \text{ fixed}},$$

always satisfies $0 \leq R(T) \leq 1$. Thus, a prior distribution could be fitted to $R(T)$ and then a change of variable

$$\theta = -T/\ell nR$$

leads to a prior distribution on θ. It is not possible to do this in this study because repeated observations with T fixed are not available from the data search. It should be noted though that if it were possible to make successive Bernoulli trials then the conditional distribution of $f(x|R)$ is a polynomial in R and the method (with the change of variable $\theta = -T/\ell nR$) could be used to fit $g(\theta)$.

Howard G. Tucker has given a method for estimating the prior distribution when the conditional distribution $f(x|\theta)$ is Poisson (An Estimate of the Compounding Distribution of a Compound Poisson Distribution, Theor. Prob. Appl. 8, 195-200, 1963). Here, the observed random variable x is not lifetime but the number of failures occurring in fixed time T. Thus, it is assumed that random samples $(x_1, \ldots, x_n)$ are available from the marginal (discrete) distribution $f(x)$. Again, the uniqueness of the moment sequence for the prior distribution is used to obtain a consistent estimate of the prior distribution.

Now, we assume throughout this study that for lifetimes, the conditional distribution of lifetime given θ is known and that though each of the n θ's is unobservable (because θ for each equipment is unknown) samples are available from $f(x|\theta)$, where X represents lifetime. The sample size from $f(x|\theta)$ will be denoted by K_i for the i^{th} θ. We are particularly concerned with methods which can deal with the conditional distribution $f_{K_i}(\hat{\theta}_{K_i}|\theta_i)$ since often the only information available to the analyst is the pair $(\hat{\theta}, K_i)$ rather than the lifetimes $(x_1, \ldots, x_{K_i})$.

The case K_i large $i=1, \ldots, n$(e.g., $K_i > 30$ all i) can be dismissed more or less out of hand since (see Section 9.5) the joint (marginal) distribution $f(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n)$ can be used directly to fit the prior distribution $g(\theta)$. Before proceeding to other cases, it is well to

---

*Roughly, a consistent estimate means that in the limit, i.e., as $n \to \infty$ the true prior distribution will be known exactly.

clear up some terminology. In reliability work the estimation of the prior distribution is called fitting a prior distribution. In some of the more theoretical work, it is called: estimating the mixing distribution or sometimes: estimating the compounding distribution. Thus, in the statistical literature, the prior distribution is sometimes called mixing distribution or the compounding distribution.

The final method we discuss for fitting an unspecified prior distribution is the Krutchkoff-Rutherford (hereafter, K-R) method. Strictly speaking, the K-R method is not usable when the prior distribution is completely unspecified. The K-R method requires that the prior distribution belong to the Pearson family of curves. This is not very restrictive since the Pearson family is a rather large family (e.g., it includes the inverted gamma distribution). It turns out that members of the Pearson family are uniquely determined by their first four moments. Thus, computations are greatly simplified (still requiring a computer program though) as regards the previous two methods discussed. As in the other methods, the K-R method leads to consistent estimates. There are two serious shortcomings with the K-R method:

i) It is not always usable.

ii) It has large and <u>unknown</u> sampling errors.

The second shortcoming is common to all the methods discussed but is particularly true of the K-R method because it uses only the first four moments. The first shortcoming above is the most serious. If the sample second central moment is negative the K-R method cannot be applied. The K-R method was tried on seventeen sets (seventeen different types of equipment) of data. Only four sets resulted in an estimate of the prior distribution, the other thirteen resulting in negative sample second moments. The K-R method is discussed in detail in Section 4.3.1 of this report.

4.2 METHODS OF FITTING PRIOR DISTRIBUTIONS WHEN THE FAMILY IS SPECIFIED

4.2.1 INTRODUCTION

The problem of fitting a prior distribution is much simpler when the prior distribution family is specified than when it is unspecified. Once a distribution family (e.g., inverted gamma), is assumed for the prior then fitting the prior is a two-fold procedure:

1) Obtaining estimates of the parameters of the prior distribution.

2) Testing the validity of the assumed prior by a goodness-of-fit test.

Before discussing how 1) and 2) are carried out the general setup is recalled:

There is a prior distribution $g(\theta)$ on the parameter θ, where θ is always considered the mean of an exponential time-to-failure distribution (i.e., $f(x|\theta) = (1/\theta)\, e^{-x/\theta}$). Sample values from some observable random variable are available. In this study, only two types of observable random variables are of sufficient practical interest to be investigated:

Type 1. The observed random variable is the number, X, of failures of a unit occurring in a fixed time T. X is a discrete variable, taking on only the values $0,1,2, \ldots$. Observations on X are obtained by putting n units on test for time T, and recording the number of failures for each unit $\{x_1, \ldots, x_n\}$.

Type 2. The observed random variable is the sample MTBF, $\hat{\theta}$. Two cases are possible here:

Type 2a. Each value of $\hat{\theta}$ is based on a fixed number of failures, K. Sample values of $\hat{\theta}$ are obtained by putting n units on test and recording K failure times $x_{i,1}, \ldots, x_{i,K}$ for each unit $i=1, \ldots, n$, and then computing

$$\hat{\theta}_i = \left(\sum_{j=1}^{K} x_{i,j}\right) \Big/ K, \quad i=1, \ldots, n.$$

(Note that since $\hat{\theta}_i$ is a sufficient statistic for $(X_{i,1}, \ldots, X_{i,K})$, one has no need of the $x_{i,j}$'s once $\hat{\theta}_i$ has been computed). $\hat{\theta}$ is a continuous random variable defined on $[0,\infty]$, (assuming, of course, that θ also is).

Type 2b. The values of $\hat{\theta}$ are based on different values of K. This is the case of practical interest. Data in the form of Case 2a does not occur in practice, but Case 2a is included because its treatment is analytically simple and sheds light on Case 2b. For this case, the observed data is in the form of n pairs $(\hat{\theta}_i, K_i)$, where K_i is the number of failure times used for the computation of $\hat{\theta}_i$.

Whether the observed data is of Type 1 or Type 2, the situation is the same: the observed data is not from the prior $g(\theta)$, but from the marginal distribution $f(x)$ (or $f(\hat{\theta})$). Hence, the data cannot be fitted "directly" to $g(\theta)$. However, when $g(\theta)$ is specified, $f(x)(f(\hat{\theta}))$ is uniquely specified, since

$$f(\bullet) = \int_0^\infty g(\theta) f(\bullet|\theta)\, d\theta$$

and $f(x|\theta)(f(\hat{\theta}|\theta))$ is specified as a particular Poisson (gamma) distribution by the exponential assumption. Since the parameters of $f(x)(f(\hat{\theta}))$ include the parameters of $g(\theta)$, one need only estimate the parameters of the marginal to obtain estimates of the prior parameters. Also, because of the unique association between the prior and the marginal, a goodness of fit test of the marginal also serves as a goodness of fit test for the prior.

In summary, a specified prior yields a specified marginal containing the same parameters. Since the observed data comes from the marginal, one can fit the data to the marginal using any appropriate classical methods of parameter estimation and goodness of fit tests. In the process, the prior $g(\theta)$ is also fitted.

In the following sections, specific formulas are derived for fitting the prior when the specified prior family is the inverted gamma distribution

$$g(\theta)=\frac{\alpha^{\lambda}}{\Gamma(\lambda)}\left(\frac{1}{\theta}\right)^{\lambda+1} e^{-(\alpha/\theta)}. \qquad (4.2.1.1)$$

The inverted gamma prior distribution is selected because it is a two parameter distribution, very flexible, and the natural conjugate when the conditional distribution is exponential. Also, when the observed data is of Type 1 or Type 2 above, the marginal density is available in closed form.

The method selected for estimating parameters is the method of matching moments. For goodness of fit, the χ^2 test is used.

4.2.2 PARAMETER ESTIMATION

4.2.2.1 PARAMETER ESTIMATION FOR TYPE 1 DATA

When the observed statistic is the number of failures, X, occurring in fixed time T, then its distribution is Poisson,

$$f(x|\theta) = \frac{e^{-T/\theta}(T/\theta)^x}{x!} \qquad T,\theta > 0 \qquad (4.2.2.1.1)$$

$$x = 0,1, \ldots$$

since the conditional distribution of time to failure for fixed θ is exponential. With a specified inverted gamma prior distribution (Formula 4.2.1.1), the marginal distribution of X, derived in Section 4.2, is given by

$$f(x) = \frac{\Gamma(\lambda+x)}{\Gamma(\lambda)x!}\left(\frac{T}{T+\alpha}\right)^x \left(\frac{\alpha}{T+\alpha}\right)^\lambda, \qquad x = 0,1, \ldots \qquad (4.2.2.1.2)$$

This is a negative binomial distribution with

$$\text{Mean} = E(X) = \frac{\lambda T}{\alpha} \qquad (4.2.2.1.3)$$

$$\text{Variance} = \sigma_x^{\ 2} = \frac{\lambda T(T+\alpha)}{\alpha^2} \qquad (4.2.2.1.4)$$

$$E(X^2) = (E(X))^2 + \sigma_x^{\ 2} = \frac{\lambda T(T+\alpha)}{\alpha^2} + \left(\frac{\lambda T}{\alpha}\right)^2 \qquad (4.2.2.1.5)$$

Suppose n sample values of X, (x_1, ..., x_n) are observed. Then by the method of matching moments, we equate

$$\left(\sum x_i\right)\Big/ n = \lambda T/\alpha \qquad (4.2.2.1.6)$$

$$\left(\sum x_i^{\ 2}\right)\Big/ n = \frac{\lambda T(T+\alpha)}{\alpha^2} + \left(\frac{\lambda T}{\alpha}\right)^2 \qquad (4.2.2.1.7)$$

The values of α and λ which satisfy Formulas 4.2.2.1.6 and 4.2.2.1.7 are the required estimates, and are denoted by $\hat{\alpha}$, $\hat{\lambda}$. The solution is

$$\hat{\alpha} = \frac{\frac{\Sigma x_i}{n}\ T}{\frac{\Sigma x_i^{\ 2}}{n} - \left(\frac{\Sigma x_i}{n}\right)^2 - \frac{\Sigma x_i}{n}} \qquad (4.2.2.1.8)$$

$$\hat{\lambda} = \frac{\hat{\alpha}\left(\frac{\Sigma x_i}{n}\right)}{T} \qquad (4.2.2.1.9)$$

These formulas will be used in Section 4.3 to obtain estimates of α and λ for eight sets of field data.

4.2.2.2 PARAMETER ESTIMATION FOR TYPE 2a DATA

Type 2a data occurs when the observed statistic is the sample MTBF, $\hat{\theta}$, where the observed number of failures is a fixed integer K. The conditional distribution $f_K(\hat{\theta}|\theta)$, is gamma, i.e.,

$$f_K(\hat{\theta}|\theta) = \left(\frac{K}{\theta}\right)^K \frac{1}{\Gamma(K)} (\hat{\theta})^{K-1} e^{-\frac{K\hat{\theta}}{\theta}} \qquad (4.2.2.2.1)$$

because of the assumption of an exponential distribution of time to failures. With a specified inverted gamma prior distribution (Formula 4.2.1.1), the marginal distribution of $\hat{\theta}$, derived in Section 9.2, is given by

$$f_K(\hat{\theta}) = \frac{\Gamma(K+\lambda)}{\Gamma(K)\Gamma(\lambda)} \left(\frac{\alpha}{\alpha+K\hat{\theta}}\right)^{\lambda} \left(\frac{K\hat{\theta}}{\alpha+K\hat{\theta}}\right)^{K-1} \left(\frac{K}{\alpha+K\hat{\theta}}\right), \quad \hat{\theta} > 0. \qquad (4.2.2.2.2)$$

This is an inverted Beta distribution with

$$E(\hat{\theta}) = \frac{\alpha}{\lambda-1}, \quad \lambda > 1 \qquad (4.2.2.2.3)$$

$$E(\hat{\theta}^2) = \frac{K+1}{K} \frac{\alpha^2}{(\lambda-1)(\lambda-2)}, \quad \lambda > 2 \qquad (4.2.2.2.4)$$

Suppose n sample values of $\hat{\theta}$, $(\hat{\theta}_1, \ldots, \hat{\theta}_n)$, are observed, each based on K failures. Then the method of matching moments leads to the pair of equations

$$\frac{\sum \hat{\theta}_i}{n} = \frac{\alpha}{\lambda-1} \qquad (4.2.2.2.5)$$

$$\frac{\sum \hat{\theta}_i^2}{n} = \frac{K+1}{K} \frac{\alpha^2}{(\lambda-1)(\lambda-2)} \qquad (4.2.2.2.6)$$

For convenience, we set $\hat{m}_1 = \frac{\Sigma\hat{\theta}_i}{n}$, $\hat{m}_2 = \frac{K}{K+1} \frac{\Sigma\hat{\theta}_i^2}{n}$

Then the solution $(\hat{\alpha},\hat{\lambda})$ is given by

$$\hat{\lambda} = \frac{2\hat{m}_2-\hat{m}_1^2}{\hat{m}_2-\hat{m}_1^2} \qquad (4.2.2.2.7)$$

$$\hat{\alpha} = \hat{m}_1(\hat{\lambda}-1) \tag{4.2.2.2.8}$$

4.2.2.3 PARAMETER ESTIMATION FOR TYPE 2b DATA

When the observed statistic is $\hat{\theta}$, but the number of failures from which $\hat{\theta}$ is computed varies from unit to unit, a modification of the method in Section 4.2.2.2 can be used to obtain estimates of α and λ.

Type 2b data is available in the form of n observations, each observation being a pair $(\hat{\theta}_i, K_i)$, $i = 1, \ldots, n$, where the positive integer K_i is the number of failures used to compute $\hat{\theta}_i$. Now, for any value of K_i, if we consider the random variable $\hat{\theta}$ based on K_i failures, we obtain from equations 4.2.2.2.3 and 4.2.2.2.4.

$$E_{K_i}(\hat{\theta}) = \frac{\alpha}{\lambda-1} \tag{4.2.2.3.1}$$

$$E_{K_i}\left(\frac{K_i}{K_i+1}\hat{\theta}^2\right) = \frac{\alpha^2}{(\lambda-1)(\lambda-2)} \tag{4.2.2.3.2}$$

Since the above formulas are valid for any value of K_i, the following procedure for estimating α and λ from the sample data $\{(\theta_i, K_i)\}_{i=1}^{n}$ seems reasonable: Solve the equations

$$\frac{\Sigma\hat{\theta}_i}{n} = \frac{\alpha}{\lambda-1} \tag{4.2.2.3.3}$$

$$\frac{\Sigma \frac{K_i}{K_i+1}\theta_i^2}{n} = \frac{\alpha^2}{(\lambda-1)(\lambda-2)} \tag{4.2.2.3.4}$$

for α and λ. For convenience, we set $\hat{m}_1 = \frac{\Sigma\hat{\theta}_i}{n}$ and

$\hat{m}_2 = \sum \frac{K_i}{K_i+1} \hat{\theta}_i^{\,2} / n$. Then the solution $(\hat{\alpha},\hat{\lambda})$ is given by

$$\hat{\lambda} = \frac{2\hat{m}_2 - \hat{m}_1^{\,2}}{\hat{m}_2 - \hat{m}_1^{\,2}} \qquad (4.2.2.3.5)$$

$$\hat{\alpha} = \hat{m}_1(\hat{\lambda}-1) \qquad (4.2.2.3.6)$$

Note, that we have obtained point estimates of α and λ without having to consider what the form of the marginal density $f(\hat{\theta})$ is when the K_i's are different. This was possible because the expectations in Formulas 4.2.2.3.1 and 4.2.2.3.2 are independent of K_i. However, when goodness of fit is discussed in the next section, it will become necessary to specify a model for the marginal distribution.

4.2.3 GOODNESS OF FIT

4.2.3.1 THE χ^2 TEST

The χ^2 test was used in this study to test the goodness of fit of the estimated prior distributions. This section will outline the steps involved in taking the χ^2 test. The theory behind the χ^2 test (and also the method of matching moments used in Section 4.2.2) can be found in many standard textbooks on Statistics.

Suppose we have specified a family of probability distributions (e.g., negative binomial of form (4.2.2.1.2)), and have used sample data to estimate q unknown parameters of the distribution. To test goodness of fit:

1) Select a significance level p (e.g., p = .90, p = .95).

2) Divide the range of the data into cells, so that at least 5 sample values lie in each of the c cells.

3) Count the number of observations in each cell, and compute the expected number of observations in each cell under the hypothesis that the data arose from the distribution with the estimated parameters.

4) Denoting the observed and expected frequencies in each cell by O_i, E_i, i=1, ..., c, compute

$$\chi^2 = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i} \qquad (4.2.3.1.1)$$

5) Compare χ^2 to the p^{th} quantile of the χ^2 distribution with c-q-1 degrees of freedom (denoted $\chi^2_{p,c-q-1}$).

If

$$\chi^2 > \chi^2_{p,c-q-1} \qquad (4.2.3.1.2)$$

reject the hypothesized distribution family (i.e., a bad fit has been demonstrated).

If

$$\chi^2 < \chi^2_{p,c-q-1}, \qquad (4.2.3.1.3)$$

then accept (i.e., good fit).

We now return to the case of fitting the inverted gamma prior for different types of marginal distributions. In each case, the χ^2 test is taken against the marginal, since that is the distribution from which the sample data arises.

4.2.3.2 THE χ^2 TEST FOR TYPE 1 DATA

When the observed statistic is the number of failures, X, occurring in fixed time T, Formula (4.2.2.1.2) is used to compute $f(i) = P(X=i)$ for $i = 0,1,2, \ldots$. For a sample of size n, the expectation of X=i is nf(i). Once the integers in the range of X are divided up into cells, the expected number of observations in each cell is easily summed.

With c cells, the χ^2 test is taken with (c-3) degrees of freedom, since 2 unknown parameters (α and λ) were estimated.

4.2.3.3 THE χ^2 TEST FOR TYPE 2a DATA

When the observed statistic is $\hat{\theta}$ based on a fixed number, K, of failures, then the marginal distribution $f_K(\hat{\theta})$ is the inverted Beta distribution given in Formula 4.2.2.2.2. For a sample size of n, the expected number of observations in a χ^2 cell with upper and lower end points, U and L, respectively, is

$$n \int_L^U f_K(\hat{\theta})\, d\hat{\theta} \qquad (4.2.3.3.1)$$

The integral is not available in closed form, but can be evaluated by making a transformation to the Beta distribution and using a table of incomplete Beta functions. If the tables do not have the appropriate parameter values for a particular case, it is always possible to use the computer techniques of either numerical integration or simulation of the distribution.

Since only 2 unknown parameters are estimated (K is known), the χ^2 test is taken with degrees of freedom 3 less than the number of cells.

4.2.3.4 THE χ^2 TEST FOR TYPE 2b DATA

When the observed data consists of n pairs $\{(\hat{\theta}_i, K_i)\}$, and $\hat{\alpha}$ and $\hat{\lambda}$ are estimated as in Section 4.2.2.3, a model for the marginal distribution $f(\hat{\theta})$ must be specified in order to take a χ^2 test. This model must account for the different K_i's on which the values of $\hat{\theta}$ are based. The most reasonable model seems to be the mixed population model, in which the marginal density takes the form

$$f_{\{K_i\}}(\hat{\theta}) = p_1 f_{K_1}(\hat{\theta}) + \ldots + p_t f_{K_t}(\hat{\theta}), \qquad (4.2.3.4.1)$$

where t is the number of distinct K_i's, and p_i is the probability that a random observation of $\hat{\theta}$ is based on K_i failures (note $\sum_{i=1}^{t} p_i = 1$). The $f_{K_i}(\hat{\theta})$'s are, of course, the inverted Beta marginal densities (Formula 4.2.2.2.2) for the case of $\hat{\theta}$ being based on K_i failures. The mixed population model assigns relative "weights" to the "single population" marginals $f_{K_i}(\hat{\theta})$, in accordance with the "prior" probabilities p_i of a sample $\hat{\theta}$ being based on K_i failures. It is easily verified that $f_{\{K_i\}}(\hat{\theta})$ is indeed a density:

$$\int_0^\infty f_{\{K_i\}}(\hat{\theta})d\hat{\theta} = \int_0^\infty \sum_{i=1}^{t} p_i f_{K_i}(\hat{\theta})d\hat{\theta} = \sum p_i \int_0^\infty f_{K_i}(\hat{\theta})d\hat{\theta}$$

$$= \sum p_i = 1. \qquad (4.2.3.4.2)$$

The model contains (t+2) unknown parameters which must be estimated before a χ^2 test can be taken: $\alpha, \lambda, p_1, \ldots, p_t$. The estimation of α and λ has already been discussed. For a sample of n $\hat{\theta}$'s, the p_i are estimated by

$$\hat{p}_i = \frac{\text{number of } \hat{\theta}\text{'s computed from } K_i \text{ failures}}{n} \tag{4.2.3.4.3}$$

for $i=1, \ldots, t$.

To compute the expected number of observations in a χ^2 cell with upper and lower end points U and L, respectively, we note that

$$n \int_L^U f_{\{K_i\}}(\hat{\theta})d\hat{\theta} = n \sum p_i \int_L^U f_{K_i}(\hat{\theta})d\hat{\theta} \ , \tag{4.2.3.4.4}$$

and the evaluation of the integrals on the right was discussed in Section 4.2.3.3.

With c χ^2 cells, the χ^2 test is taken with c-t-3 degrees of freedom, since t+2 parameters of $f_{\{K_i\}}(\hat{\theta})$ are estimated.

4.3 RESULTS AND DEVELOPMENT OF PRIOR DISTRIBUTIONS FOR DATA COLLECTED

4.3.1 RESULTS OF FITTING THE PRIOR DISTRIBUTION WHEN THE FAMILY IS UNSPECIFIED

The methods for fitting the prior distribution when the family is unspecified were discussed in Section 4.1. Only one of these methods was used for fitting actual field data in this study: the Krutchkoff-Rutherford method. The necessary type of data for using the K-R method (pairs $(\hat{\theta}_i, K_i)$) was available for all 17 sets of data collected (See Section 2). Table 4.3.1 summarizes the results of applying the K-R method to the 17 sets of data (numbered as in Section 2). The formulas at the right are those used in calculations. The $\hat{m}_i$'s are the unbiased estimates of the moments of $g(\theta)$ that the K-R method calls for. The $\hat{\mu}_i$'s are estimates of the central moments of $g(\theta)$ obtained by using the relationships between the central moments and the moments about 0 that hold for any distribution. Finally, the estimates of the μ_i are used to obtain estimates of β_1, β_2 and κ, which are used in accordance with criteria in Elderton's book (Reference 2) to identify uniquely the appropriate Pearson curve type.

In Table 4.3.1, Pearson curve types are identified for only Data Sets Nos. 2, 11, 12, and 14, which are fitted as Types VI, I, I, and IV respectively. For the 13 other cases, the K-R method fails to work, since the estimates of μ_2 are negative. The failure to fit a prior in 13 out of 17 cases, is, of course, no fault of the data, but is inherent in the K-R method. Even in the four cases in which a prior distribution is identified, the results

Table 4.3.1 - Results of Data Analysis by the [illegible] - Rutherford Method

Data Set No.	$\hat{\nu}_1$	$\hat{\nu}_2$	$\hat{\nu}_3$	$\hat{\nu}_4$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\beta}_1$	$\sqrt{\hat{\beta}_1}$	a	k	Pearson Type	Formulas
1	[illegible]	3.93×10^{6}	5.25×10^{9}	5.47×10^{12}	-9.09×10^{5}	6.03×10^{8}	3.15×10^{12}	*	*	*	*		$\hat{\nu}_1 = \left(\sum_{i=1}^{n} \hat{\theta}_i\right)/n$
2	1508	2.29×10^{6}	2.88×10^{9}	2.87×10^{12}	3.72×10^{6}	-7.00×10^{8}	1.48×10^{12}	9549	97.72	2892	4.275	VI	
3	2013	3.35×10^{6}	4.27×10^{9}	4.38×10^{12}	-7.03×10^{5}	3.55×10^{8}	2.11×10^{12}	*	*	*	*		$\hat{\nu}_2 = \left(\sum_{i=1}^{n} \frac{K_i}{K_i+1} \hat{\theta}_i^2\right)/n$
4	7404	1.38×10^{6}	5.74×10^{9}	5.90×10^{12}	-1.46×10^{6}	2.36×10^{9}	3.97×10^{11}	*	*	*	*		
5	3013	5.90×10^{6}	8.15×10^{9}	8.61×10^{12}	-3.18×10^{6}	9.56×10^{9}	-1.57×10^{13}	*	*	*	*		$\hat{\nu}_3 = \left(\sum_{i=1}^{n} \frac{K_i^2}{(K_i+1)(K_i+2)} \hat{\theta}_i^3\right)/n$
6	2411	4.40×10^{6}	5.98×10^{9}	6.26×10^{12}	-1.41×10^{6}	2.12×10^{9}	1.25×10^{12}	*	*	*	*		
7	2007	3.71×10^{6}	4.90×10^{9}	5.13×10^{12}	-6.48×10^{5}	-9.71×10^{7}	3.06×10^{12}	*	*	*	*		$\hat{\nu}_4 = \left(\sum_{i=1}^{n} \frac{K_i^3}{(K_i+1)(K_i+2)(K_i+3)} \hat{\theta}_i^4\right)/n$
8	2067	4.00×10^{6}	5.27×10^{9}	5.48×10^{12}	-1.14×10^{6}	1.35×10^{9}	1.80×10^{12}	*	*	*	*		
9	112	1.13×10^{4}	1.54×10^{6}	2.40×10^{8}	-1.25×10^{3}	5.52×10^{5}	-7.14×10^{7}	*	*	*	*		
10	383	9.84×10^{4}	1.62×10^{7}	2.28×10^{9}	-5.40×10^{4}	2.22×10^{7}	-5.74×10^{9}	*	*	*	*		$\hat{\mu}_2 = \hat{\nu}_2 - \hat{\nu}_1^2$
11	862	2.56×10^{6}	9.49×10^{9}	2.84×10^{13}	1.80×10^{6}	3.09×10^{9}	6.40×10^{12}	2.579	1.006	1.97	-11.5	I	$\hat{\mu}_3 = \hat{\nu}_3 - 3\hat{\nu}_1\hat{\nu}_2 + 2\hat{\nu}_1^3$
12	885	2.52×10^{5}	2.26×10^{8}	1.93×10^{11}	1.72×10^{5}	5.76×10^{7}	3.88×10^{10}	.6554	.8096	1.308	-.1739	I	
13	72	3.63×10^{3}	1.1×10^{5}	4.41×10^{6}	-1.50×10^{3}	9.63×10^{4}	-3.25×10^{6}	*	*	*	*		$\hat{\mu}_4 = \hat{\nu}_4 - 4\hat{\nu}_1\hat{\nu}_3 + 6\hat{\nu}_1^2\hat{\nu}_2 - 3\hat{\nu}_1^4$
14	438	7.04×10^{5}	1.83×10^{9}	7.58×10^{12}	5.52×10^{5}	4.21×10^{8}	1.75×10^{11}	1.053	1.026	.5768	.4886	IV	
15	2535	6.14×10^{6}	1.13×10^{10}	1.67×10^{13}	-3.86×10^{5}	-2.77×10^{9}	1.42×10^{13}	*	*	*	*		$\hat{\beta}_1 = \hat{\mu}_3^2/\hat{\mu}_2^3$
16	50	2.46×10^{3}	8.56×10^{4}	2.38×10^{6}	-9.93×10^{2}	5.55×10^{4}	-2.05×10^{6}	*	*	*	*		$\hat{\beta}_2 = \hat{\mu}_4/\hat{\mu}_2^2$
17	108	6.78×10^{3}	3.13×10^{5}	1.20×10^{7}	-4.95×10^{3}	6.46×10^{5}	-5.84×10^{7}	*	*	*	*		

$$\hat{k} = \frac{\hat{\beta}_1(\hat{\beta}_2+3)^2}{4(4\hat{\beta}_2 - 3\hat{\beta}_1)(2\hat{\beta}_2 - 3\hat{\beta}_1 - 6)}$$

$$\left\{\begin{array}{l}\hat{\lambda} = (2\hat{\nu}_2 - \hat{\nu}_1^2)\ (\hat{\nu}_2 - \hat{\nu}_1^2) \\ a = \hat{\nu}_1(\hat{\lambda} - 1)\end{array}\right.$$

obtained by solving

$$\left\{\begin{array}{l}\hat{\nu}_1 = a/(\hat{\lambda} - 1) \\ \hat{\nu}_2 = a^2/((\hat{\lambda} - 1)(\hat{\lambda} - 2))\end{array}\right.$$

* NOTE: Whenever $\hat{\nu}_2 - \hat{\nu}_3 - \hat{\nu}_1^2 < 0$, further calculations make no sense and are omitted.

are not very credible because of the large and unknown sampling errors inherent in the method. To see just how "bad" the errors are, some simulations of the K-R method (to be discussed in Section 5) were carried out. The results indicate that the K-R method gives hopelessly bad results when used on data of any practical sample size.

A list of the computer program used for the K-R method is given in Section 10.

4.3.2 RESULTS OF FITTING THE PRIOR DISTRIBUTION WHEN THE FAMILY IS SPECIFIED

The results of fitting an inverted gamma prior distribution are shown in Table 4.3.2.1. The 8 sets of field data used are those given in Table 2.3.2.1. In all 8 sets of data, all units were put on test for time T = 4320 hours and the number of failures, X, recorded. (None of the other 9 sets of data were based on fixed-time testing). Therefore, the method of parameter estimation used is that given in Section 4.2.2.1 (Type 1 data). Table 4.3.2.1 gives, for each set of data, the sample mean $\Sigma x_i/n$, the sample variance $\Sigma x_i^2/n - (\Sigma x_i/n)^2$, and the inverted gamma parameter estimates $\hat{\alpha}$ and $\hat{\lambda}$, calculated by Formulas 4.2.2.1.8 and 4.2.2.1.9, respectively. The last 4 columns in the table pertain to the χ^2 goodness-of-fit tests as described in Sections 4.2.3.1 and 4.2.3.2. The number, c, of χ^2 cells selected is given, along with the computed value of χ^2, and a designation of whether the χ^2 test passes or fails (with c-3 degrees of freedom) at both significance levels p = .99 and p = .90.

From Table 4.3.2.1, one can see that at the .99 significance level, the χ^2 test is passed in 7 out of 8 cases. Since each case is for a different type of equipment, these results indicate a general applicability of the inverted gamma prior distribution on θ for a large range of equipment types. Figures 4.3.2.1 through 4.3.2.14 give the empirical and theoretical marginal distributions and the theoretical inverted gamma prior distributions which have been fitted.

In Section 10, there is a list of the computer program used to calculate $\hat{\alpha}$, $\hat{\lambda}$, and χ^2 from the field data. Included is a list of the raw data for each of the 8 cases, indicating the grouping into χ^2 cells. A listing of the output of the program is also included.

An attempt was made to fit a set of field data to a Weibull prior distribution:

$$g(\theta) = \frac{\beta}{\alpha}\,\theta^{\beta-1} e^{-\theta^{\beta}/\alpha}. \qquad (4.3.2.1)$$

Data Set No. 4, with estimated inverted gamma prior parameters $\hat{\alpha}$ = 3609.6 and $\hat{\lambda}$ = 2.605, was selected. Weibull parameters were selected to yield the same mean and variance as the above inverted gamma distribution. The estimates of the Weibull scale and shape parameters are then, respectively,

TABLE 4.3.2.1 FIELD DATA FITTED TO INVERTED GAMMA PRIOR DISTRIBUTIONS

Data Set No.	Sample Mean	Sample Variance	Parameter Estimates $\hat{a}$	$\hat{\lambda}$	No. Of χ^2 Cells	χ^2	Pass At .99 Level	Pass At .90 Level
1	4.22	18.5	1279.67	1.24991	4	1.35756	Yes	Yes
2	6.42	43.0	758.591	1.12716	8	6.07651	Yes	Yes
3	4.49	31.2	727.436	.75632	6	21.9927	No	No
4	3.12	6.85	3609.57	2.60495	4	1.31484	Yes	Yes
5	2.3	3.61	7584.73	4.03817	4	5.60818	Yes	No
6	3.33	9.71	2251.61	1.73419	5	1.50224	Yes	Yes
7	4.59	15.8	1767.69	1.87662	5	5.64293	Yes	Yes
8	3.88	25.6	770.804	.692296	4	7.67972	Yes	Nc

$\hat{\alpha} = 385.65$, $\hat{\beta} = .7856$. In Table 4.3.2.2, the χ^2 cells for Data Set No. 4 are shown, giving for each cell the observed and the expected number of observations under both the inverted gamma hypothesis and the Weibull hypothesis. The expected values for the inverted gamma hypothesis are computed as described in Section 4.2.3.2. For the Weibull case, however, the marginal distribution is not available in closed form, so it had to be simulated. For $T = 4320$, $\alpha = 385.65$, and $\beta = .7856$, 10,000 random values of X (number of failures in time T) were drawn in a two-stage process: first, a random θ is drawn from the Weibull prior, then a random number of failure times in T given θ is drawn. The relative frequencies in 10,000 trials of the events $\{X=0, X=1, \ldots\}$ are used to approximate the marginal distribution $f(x)$.

The calculated values of χ^2 are 1.315 and 16.482 for the inverted gamma hypothesis and the Weibull hypothesis, respectively. Since in each case two unknown parameters are estimated, both χ^2 tests are taken with $4-1-2 = 1$ degree of freedom. For the inverted gamma case, the χ^2 test is passed at significance level $p = .90$, and even with p as low as it is passed after .70. The Weibull hypothesis, however, fails the χ^2 test even at level $p = .9999$. (Note that p is the probability of acceptance when the hypothesis is true, and thus it is easier to pass a χ^2 test when p is high). Since the hypothesized Weibull prior fails the χ^2 test so decisively, this result significantly helps to validate the assumption of an inverted gamma prior.

In this study no type 2a data was uncovered although by a designed test such data certainly could be gathered. Much of the data gathered in this study was of type 2b; in fact, seventeen (17) sets. Thus, the mixed marginal distribution model had to be applied. This model was discussed in Subsection 4.2.3.4. In order to apply the χ^2 test for type 2b it is necessary to have the number of cells in the χ^2 test, c, such that $c \geq t+4$. In none of the seventeen (17) data sets was the sample size n large enough so that at least t+4 cells were available. This is not as severe a limitation as it first might appear. The number of identical equipments n should be as large as possible for good fits and if n is relatively large with respect to the number of distinct K's, i.e., t the mixed model can be applied.

4.3.3 REMARKS ON THE FITTED PRIOR DISTRIBUTIONS

Data sets 1, 2, 4, 5, 6, 7, 8 have been shown to be well described by an inverted gamma prior distribution. These seven data sets represent seven different types of equipment. The question arises as to the degree of applicability of these results to other similar equipments. For example, Data set 4 represented an oscilloscope. There are many different types of oscilloscopes

TABLE 4.3.2.2. χ^2 TESTS FOR INVERTED GAMMA AND WEIBULL PRIOR DISTRIBUTIONS

- $n = 51$
- $T = 4320$
- Inverted Gamma parameters $\hat{\alpha} = 3609.6$, $\hat{\lambda} = 2.605$
- Weibull parameters $\hat{\alpha} = 385.65$, $\hat{\beta} = .7856$

χ^2 Cell No.	Values of X in χ^2 Cell	Observed	Expected Under Inverted Gamma Hypothesis	Expected Under Weibull Hypothesis
1	0, 1	18	15.881	14.137
2	2	11	9.148	5.697
3	3, 4	12	13.491	7.543
4	5+	10	12.480	23.623

$\chi^2 = 1.3148$ for Inverted Gamma

$\chi^2 = 16.482$ for Weibull

by various manufacturers and one may ask: would these oscilloscopes have the same prior distribution as that of Data set 4? (and hence, the same estimated prior distribution as we have achieved could be used). The answer is strictly speaking, no. However, this is probably speaking too strictly for practical purposes. It is very likely that equipments similar to those used in this report will have prior distributions of the same family (i.e., inverted gamma) and hence, one need only estimate the parameters and not go through the sometimes tedious process of fitting the prior distribution itself It may also turn out that even if the prior distributions do differ, they may be combined into one. This is actually not recommended because the combination of two inverted gamma distributions in this way leads to a mixed distribution which is not inverted gamma. Clearly, there are two important areas for future study

i) The applicability of fitted prior distributions to similar equipments.

ii) The feasibility of combining equipments into one prior distribution.

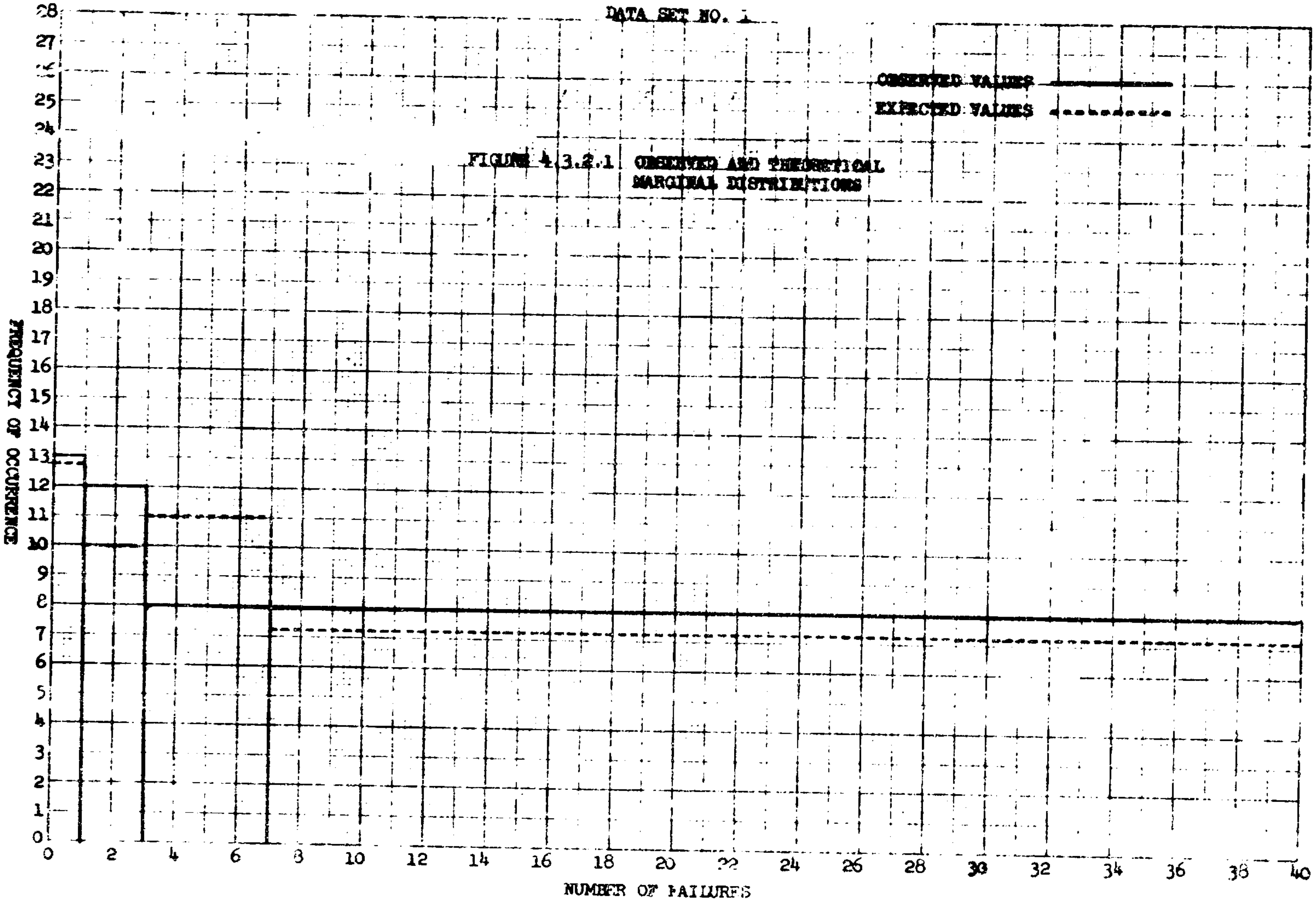
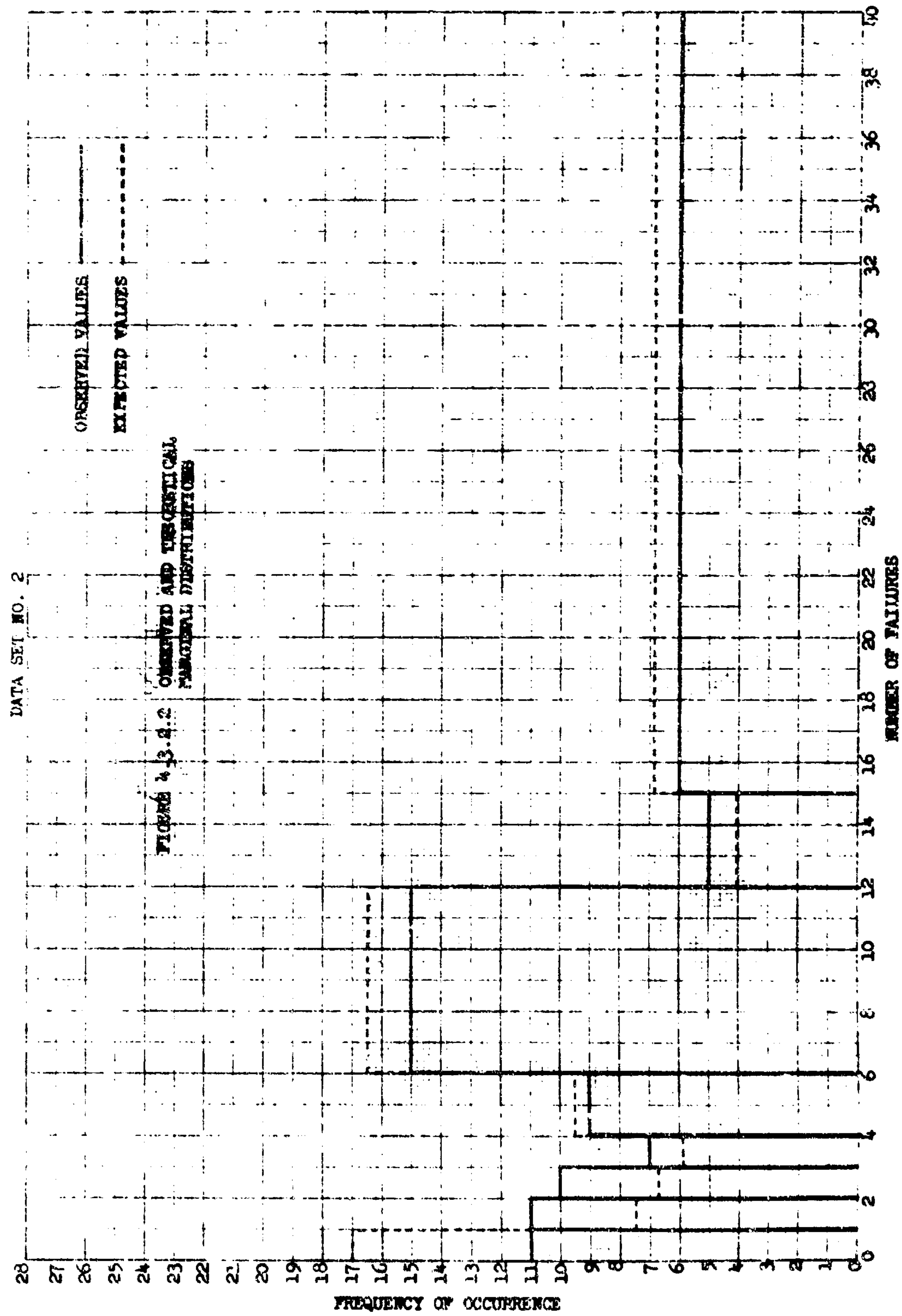


FIGURE 4.3.2.1 OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS



DATA SET NO. 2
FIGURE 4.3.2.2 OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS
OBSERVED VALUES
EXPECTED VALUES
NUMBER OF FAILURES
FREQUENCY OF OCCURRENCE

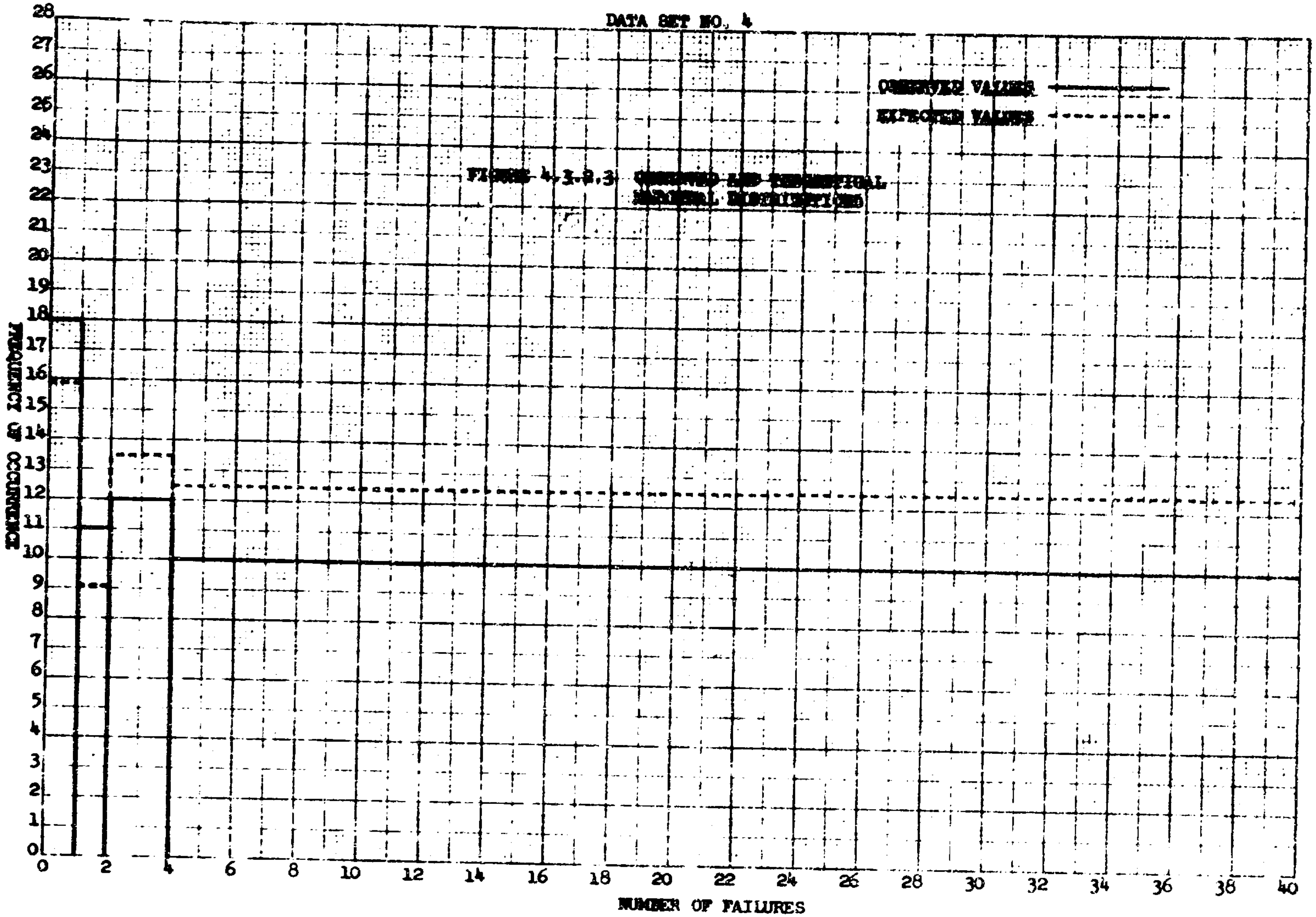


FIGURE 4.3.2.3 OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS

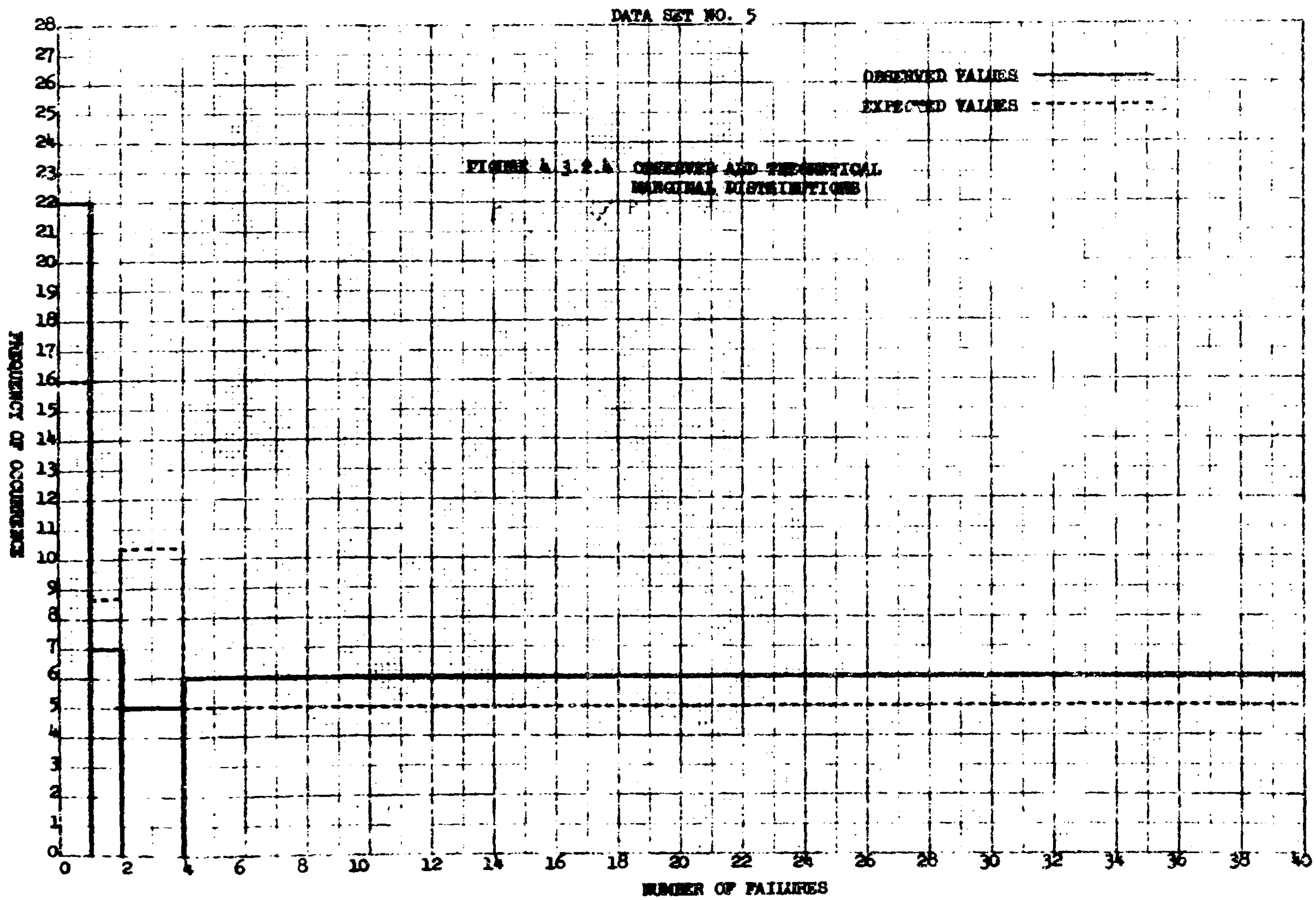
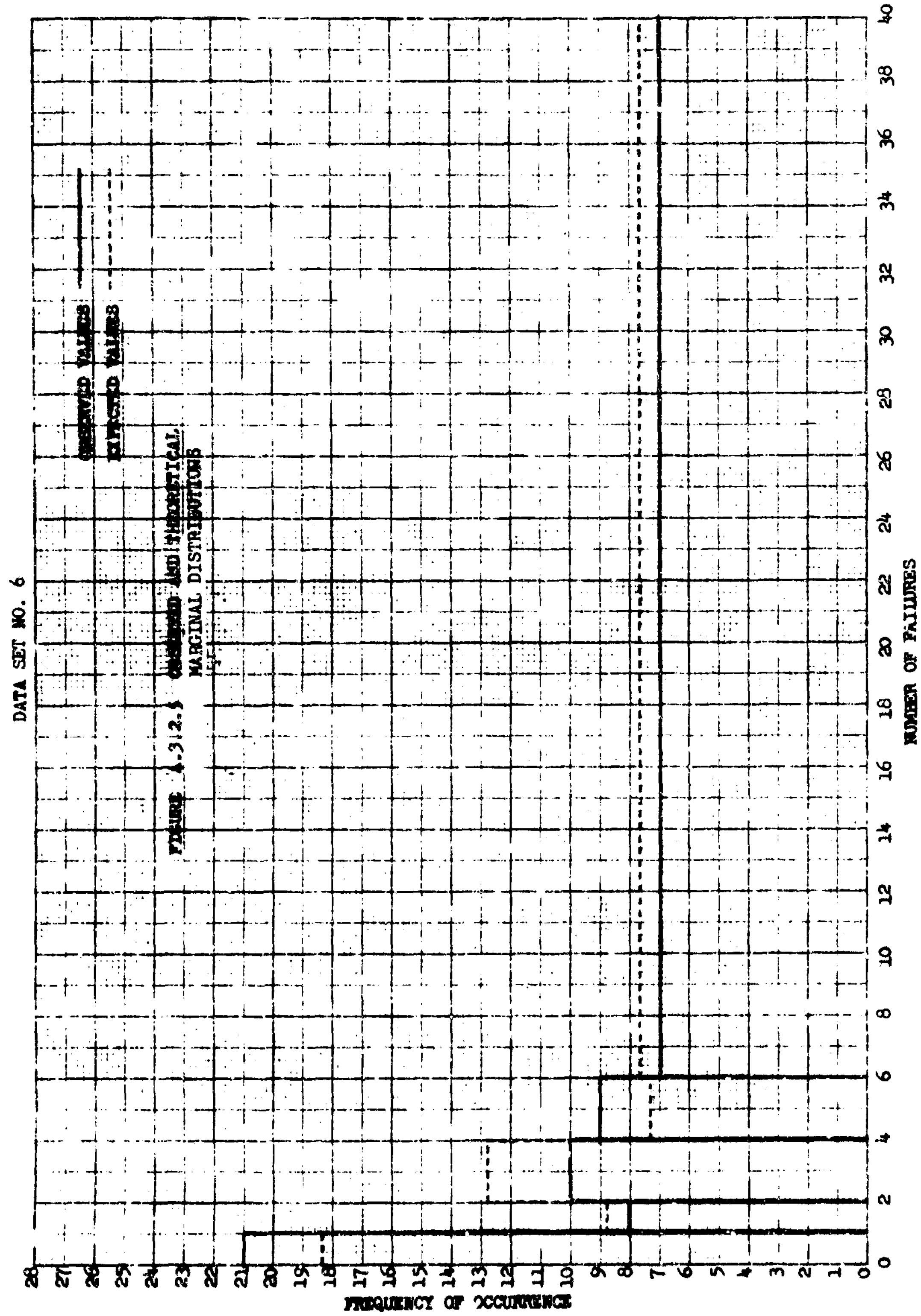
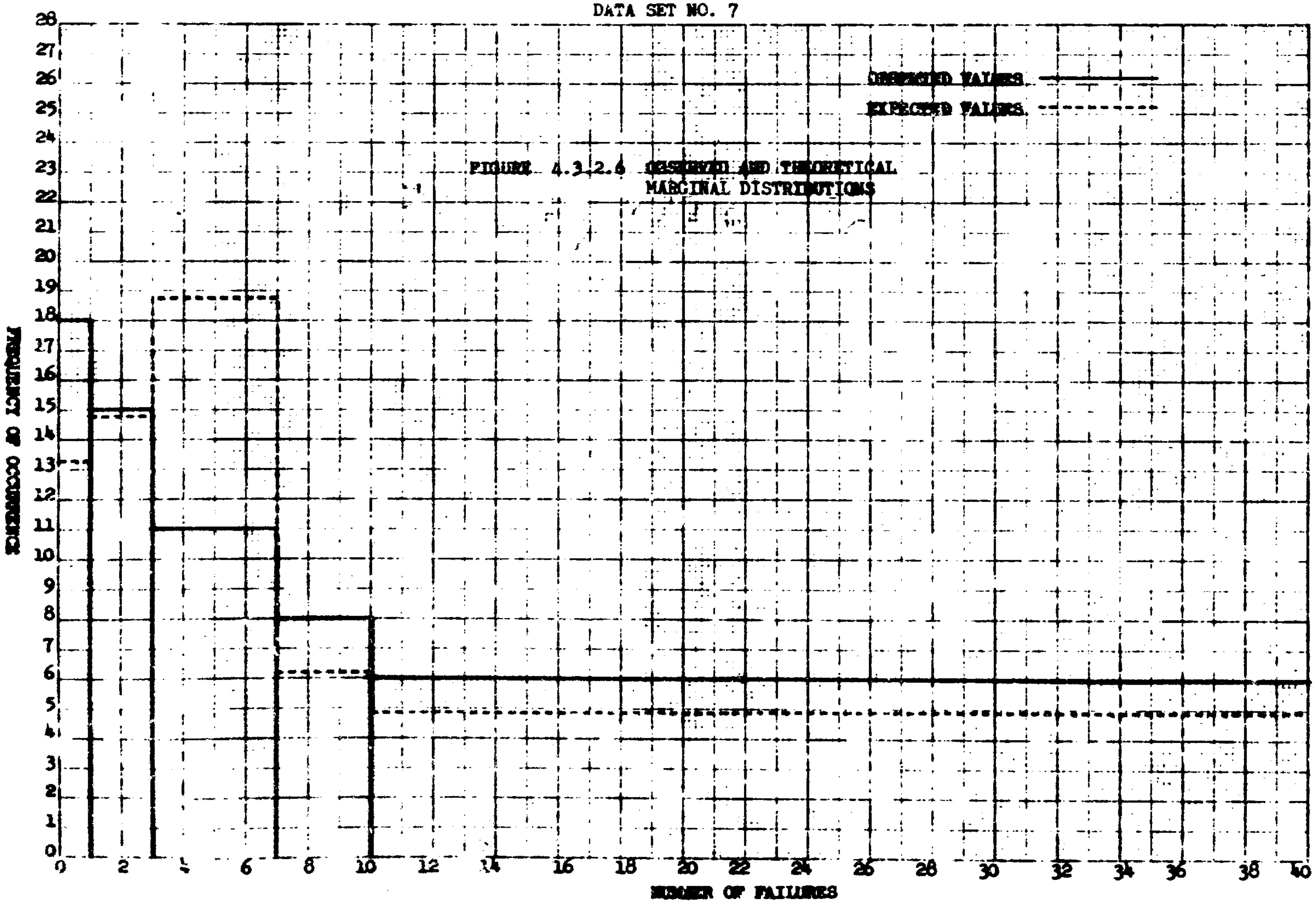


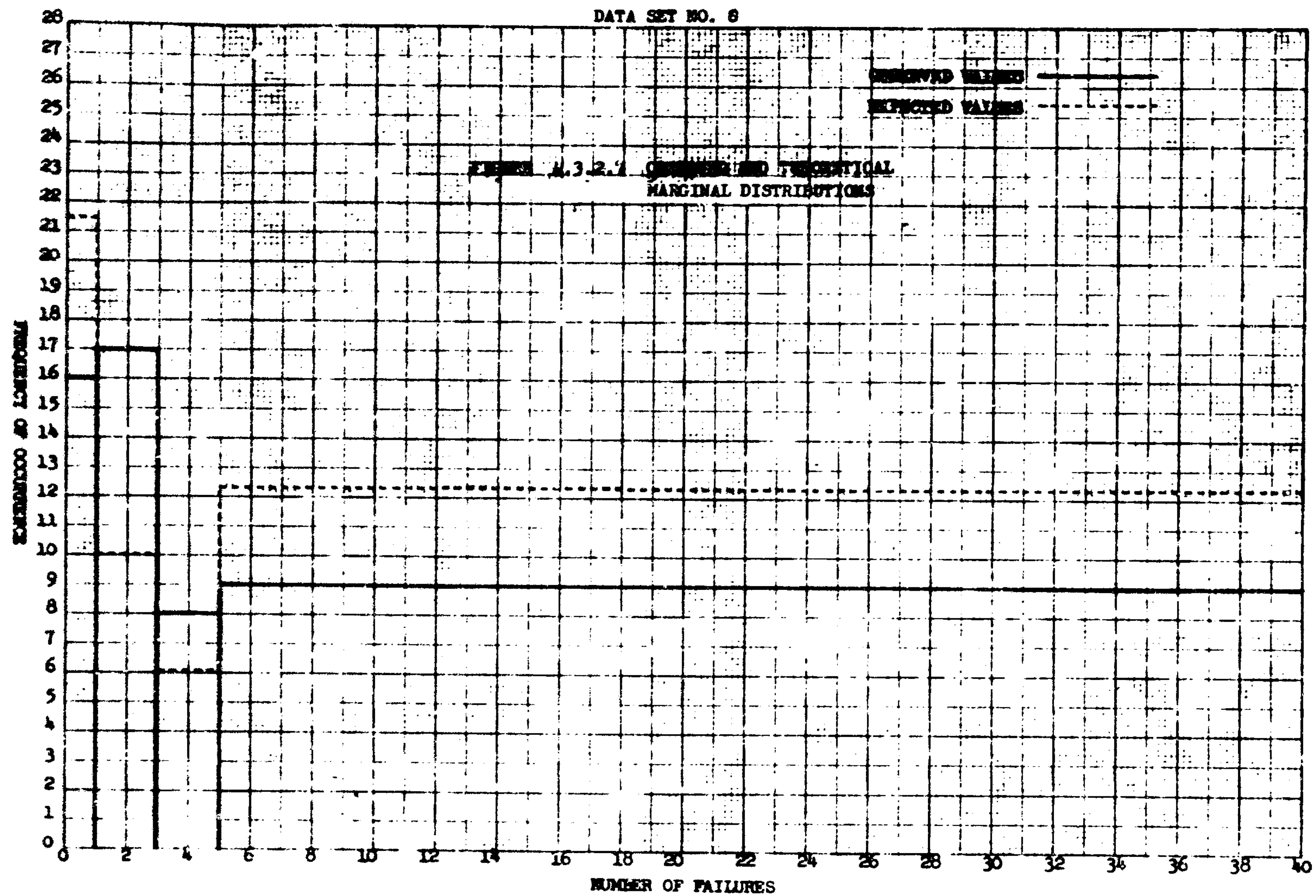
FIGURE A.3.2.A OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS


DATA SET NO. 6
FIGURE 4.3.2.5 OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS
OBSERVED VALUES
EXPECTED VALUES
NUMBER OF FAILURES
FREQUENCY OF OCCURRENCE

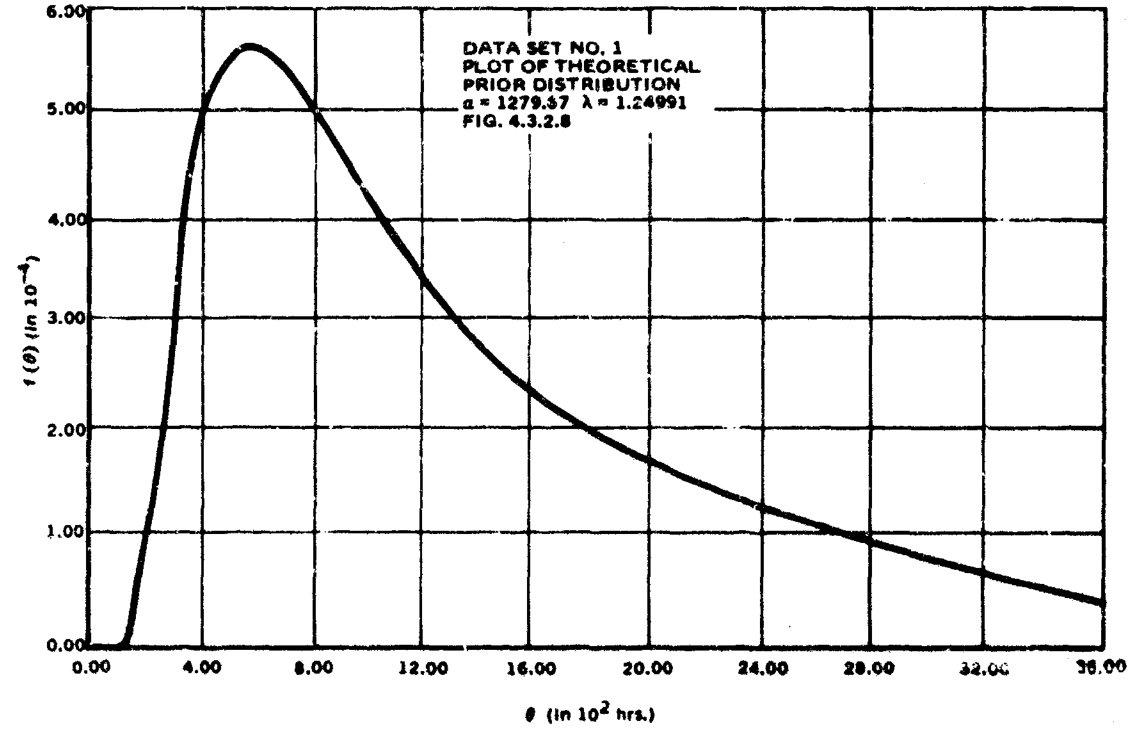
DATA SET NO. 7

FIGURE 4.3.2.6 OBSERVED AND THEORETICAL MARGINAL DISTRIBUTIONS

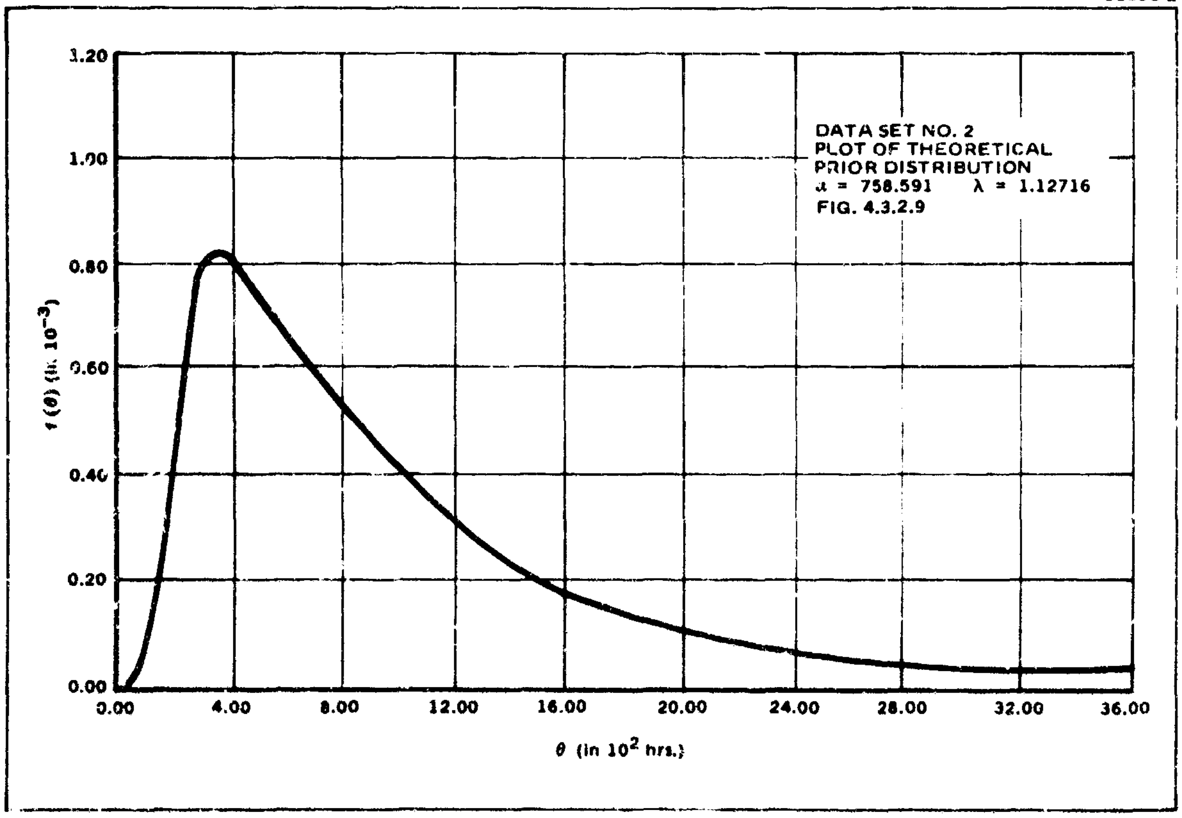


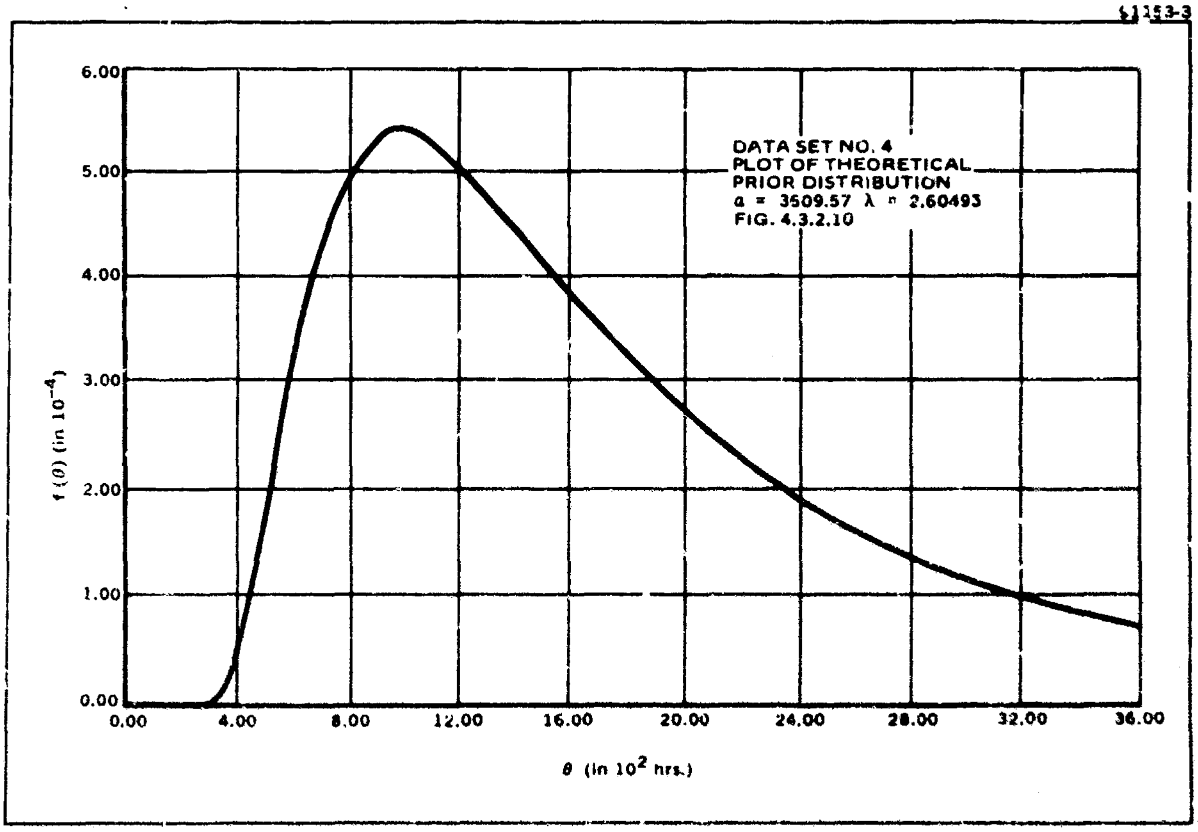

DATA SET NO. 6
MARGINAL DISTRIBUTIONS
FREQUENCY OF OCCURRENCE
NUMBER OF FAILURES
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40

91153-1

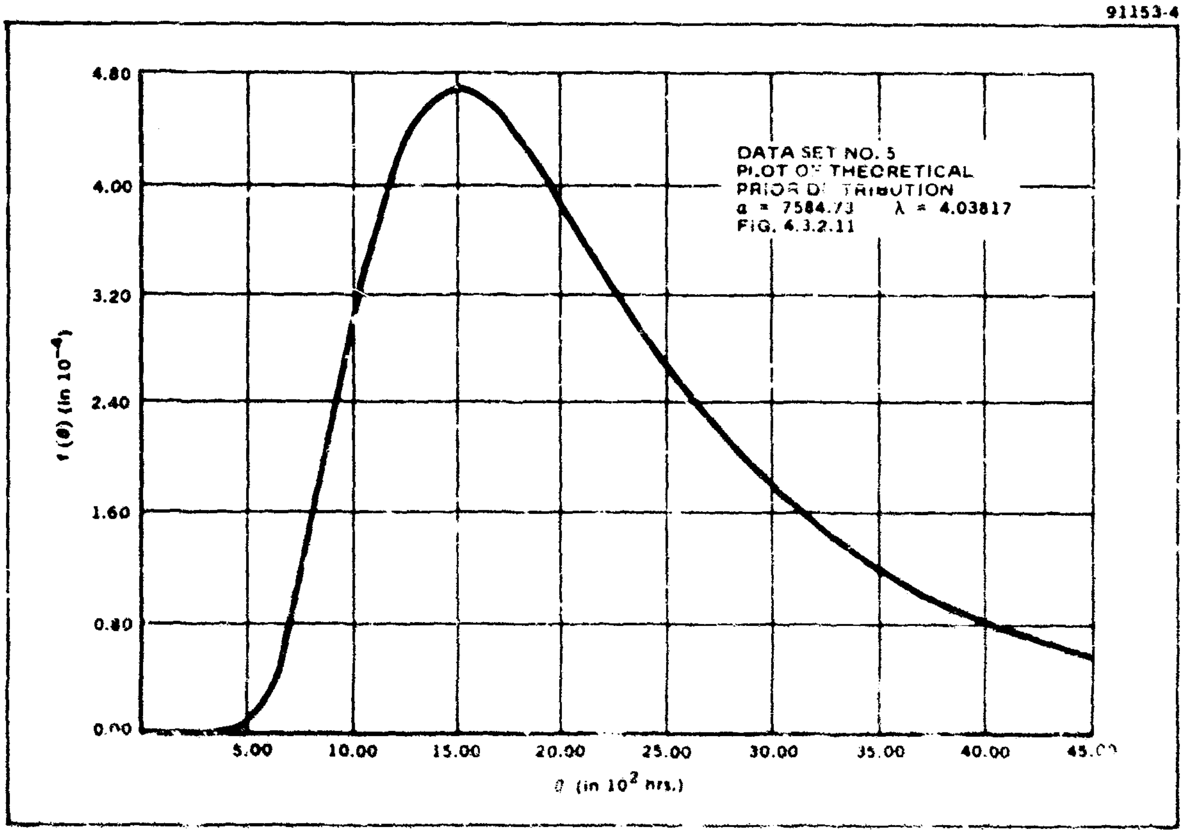

DATA SET NO. 1
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
α = 1279.57 λ = 1.24991
FIG. 4.3.2.8
f (θ) (in 10⁻⁴)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
θ (in 10² hrs.)

91153-2

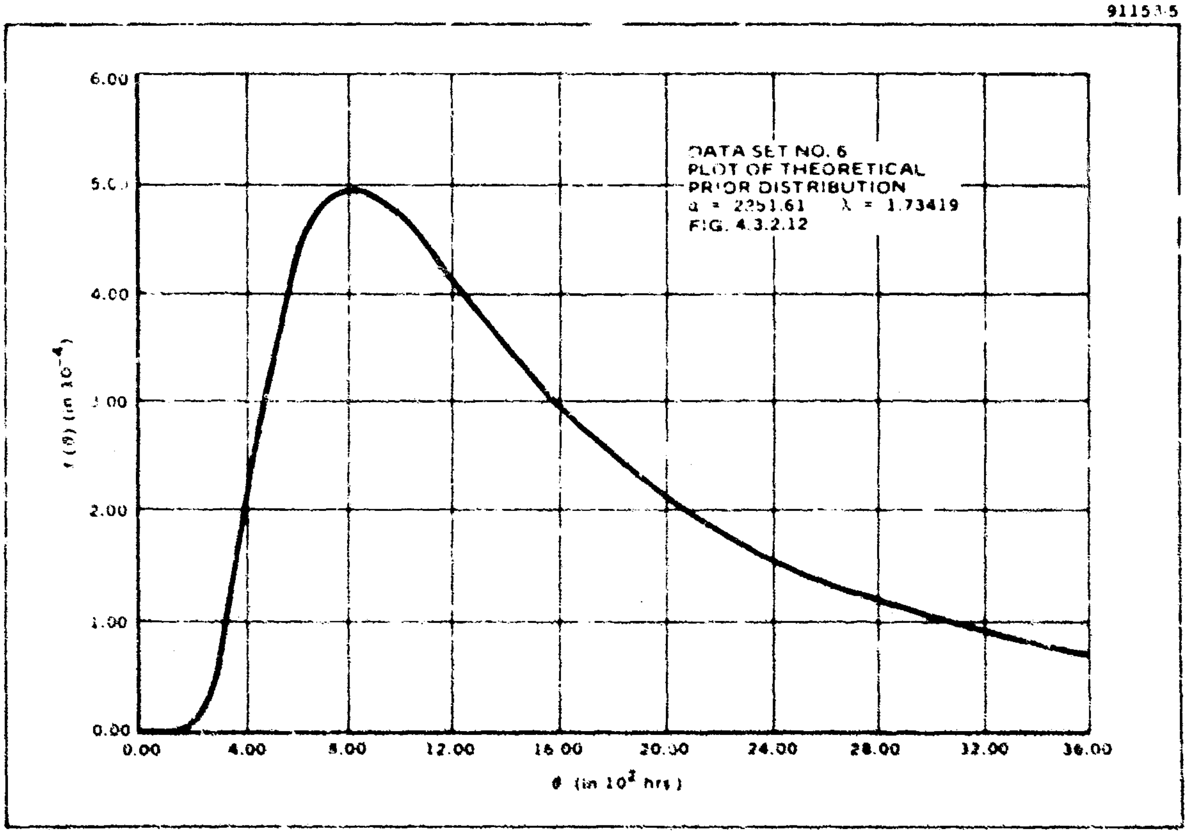

DATA SET NO. 2
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
α = 758.591 λ = 1.12716
FIG. 4.3.2.9
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
θ (in 10² hrs.)


61153-3
DATA SET NO. 4
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
a = 3509.57 λ = 2.60493
FIG. 4.3.2.10
6.00
5.00
4.00
3.00
2.00
1.00
0.00
f(θ) (in 10⁻⁴)
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
θ (in 10² hrs.)

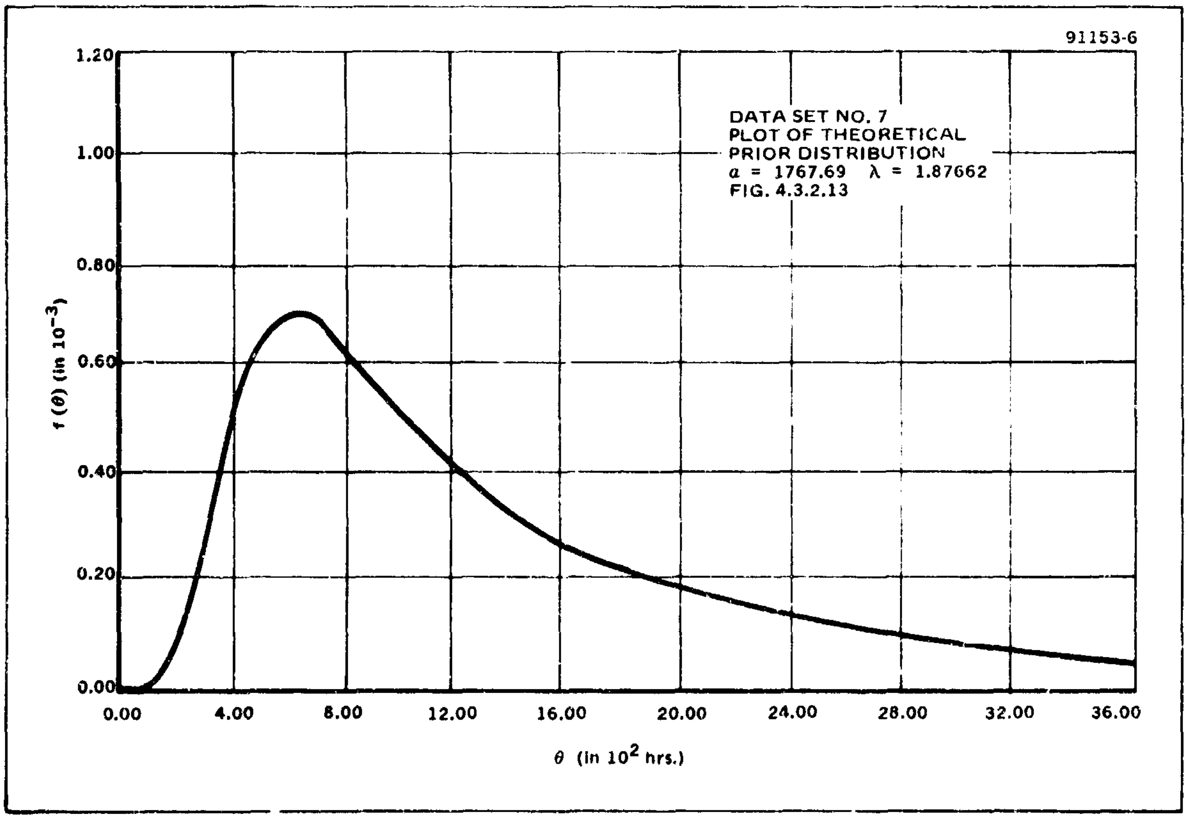
91153-4


DATA SET NO. 5
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
α = 7584.73 λ = 4.03817
FIG. 4.3.2.11
f (θ) (in 10⁻⁴)
4.80
4.00
3.20
2.40
1.60
0.80
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
θ (in 10² hrs.)

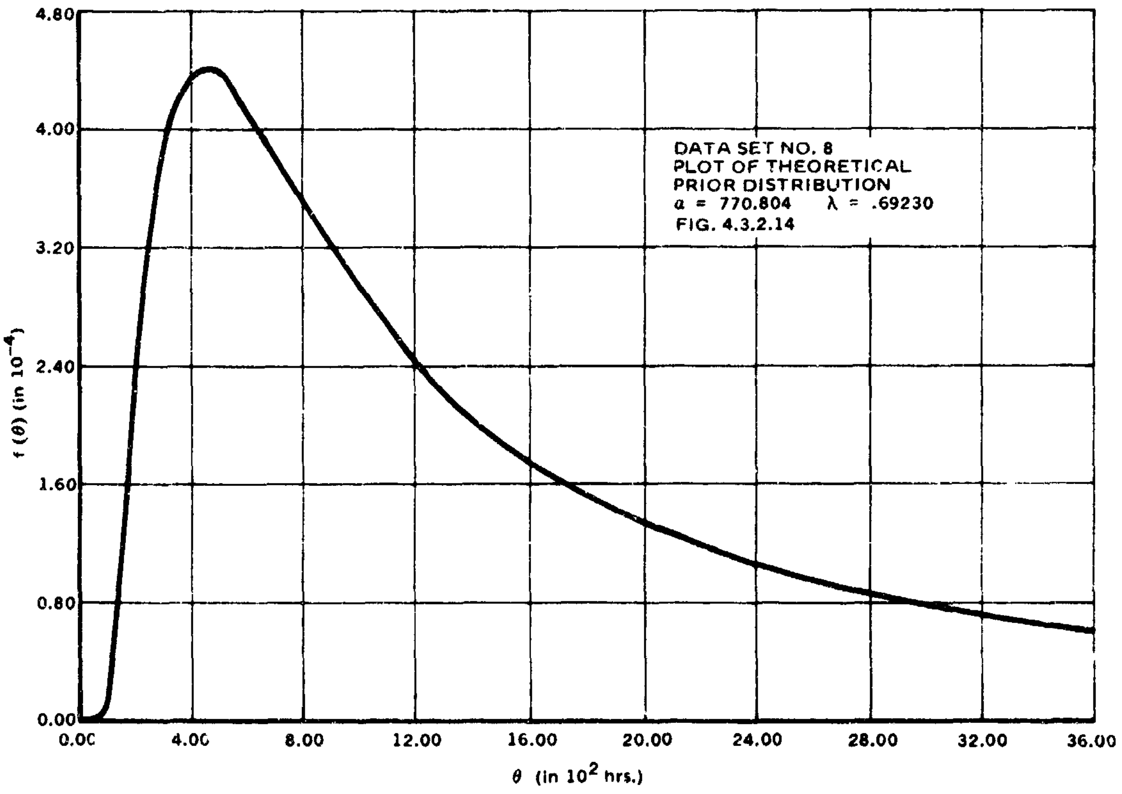
911535



DATA SET NO. 6
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
FIG. 4.3.2.12
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00


91153-6
DATA SET NO. 7
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
a = 1767.69 λ = 1.87662
FIG. 4.3.2.13
1.20
1.00
0.80
0.60
0.40
0.20
0.00
f (θ) (in 10^{-3})
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
θ (in 10^2 hrs.)

91153-7



DATA SET NO. 8
PLOT OF THEORETICAL
PRIOR DISTRIBUTION
α = 770.804 λ = .69230
FIG. 4.3.2.14

SECTION 5.0 REQUIREMENTS FOR SUITABLE A PRIORI DATA

SUMMARY

Objective

The objective of this section is to develop requirements, qualitative and quantitative, on the type and amount of data suitable for fitting a prior distribution to θ = MTBF.

Results

Regarding type of data, it was found that virtually all types of failure data are suitable if there is enough of it. Even observations on probability of survival can be used. The "preferred" types are the aforementioned type 1 and type 2a but type 2b can also be used. The results for the amount of data are summarized below.

Type 1 - number of failures in fixed time T.

$n \geq 30$, T large enough to obtain variation in the number of failures observed on each equipment. For example, T large enough so that the number of failures are not all the same for all n equipments.

Type 2a - observed MTBF, $\hat{\theta}$, K identical for all n equipments.

$n > 20$, $K > 10$.

Type 2b - observed MTBF, $\hat{\theta}$, K not the same for all equipments.

The number of identical equipments n should be large enough so that $n \geq 5t + 20$ where t is the number of distinct K_i's.

Along with the development of methods of fitting prior distributions, an important goal of this study is to establish minimal data requirements for using these methods. There are two aspects of the suitability of data to be considered:

i) The type of data.

ii) The amount of data.

5.1 THE TYPE OF DATA

Failure data can occur in several forms. We divide these into two classes: attributes and variables data. The attributes situation occurs when an equipment is operated and its survival or nonsurvival for time T (usually mission time) is observed. Then usually the binomial distribution describes such observations and the parameter probability of survival in the binomial distribution is the reliability function

$$R(T) = e^{-T/\theta}, \text{ T fixed.} \qquad (5.1.1)$$

Now, since the MTBF = θ is being considered a random variable in this study, then so is R(T) and either Rolph's method or the K-R method (See Section 4.1) can be used to fit a prior distribution to R. The change of variable $\theta = -T/\ln R$ then leads to a prior distribution on θ. For that matter, since the posterior distribution is eventually of interest the posterior distribution of θ can be obtained by using the above change of variable in the posterior distribution of R. Another attributes situation arises when test time T is fixed and the equipment is immediately repaired when a failure occurs. Then (because of the exponential assumption) the number of failures occurring in time T is Poisson and the K-R or Tucker method may be used to fit $g(\theta)$.

The variables data situation occurs when the actual failure times are available. These times occur either by agreeing to stop testing after either a fixed number of failures have occurred or after a fixed time has elapsed. The former case is called a censored test and the latter is called a truncated test. Often the failure times themselves are not available but the pair ($\hat{\theta}$ = observed MTBF, K = number of failures) is available. It really doesn't matter for all members of the class of sufficient statistics for θ (to which $\hat{\theta}$ and the failure times belong) result in the same posterior distribution for θ. Under the exponential assumption the distribution of $\hat{\theta}$ is gamma and the K-R method can be used.

Thus, virtually all forms of failure data can be used to fit prior distributions. However, in this study, only the following two types of variables data were found to be of practical interest:

Type 1 - the observed data is X, the number of failures in a fixed time T.

Type 2 - the observed data is $\hat{\theta}$, the sample mean.

Both Type 1 and Type 2 data were used in this study, with the methods of fitting priors and the results given in Section 4 of this report. The discussion of quantitative data requirements in the next section is restricted to just these two types of data.

5.2 THE AMOUNT OF DATA

In order to establish quantitative requirements for fitting a prior distribution, the following information is necessary:

1) The type of data used.

2) The method used to fit the prior distribution.

3) The criteria established to define "suitable" results for each method and type of data.

Because of the dependence of the data requirements on the above, each of the cases of Section 4 must be analyzed separately. For each case (i.e., method and data type), a sensitivity analysis will be carried out, and data requirements will be set by applying suitability criteria to the results of the analysis.

5.2.1 THE AMOUNT OF DATA-FAMILY UNSPECIFIED

The only method used in this study for fitting the prior distribution when the family is unspecified is the Krutchkoff-Rutherford method. The K-R method was used on field data in this study, and the results are summarized in Section 4.3.1. The purpose of this section is to derive suitable quantitative data requirements for applying the K-R method.

The type of data necessary for using the K-R method is observations $(\hat{\theta}_i, K_i)$ on n units. Since the data is of Type 2, it is necessary to set minimal quantitative requirements for the following:

1) the unit sample size, n.

2) the number of failures, $\{K_i\}$, i=1, ..., n.

For the purpose of determining these data requirements, the following computer simulation was carried out:

A prior distribution $g(\theta)$ was assumed to be inverted gamma with parameters $\alpha = 1500$, $\lambda = 5$. From Section 9.1, it can be seen that the condition $\lambda > 4$ guarantees that the first four moments of $g(\theta)$ are finite. Forty-two pairs (n,K) were selected, by combining 6 values of n (n = 10,20,30,50,100, 200) with each of 7 values of K(K = 5,10,20,30,50,100,200). For each (n,K) pair, n random values of $\hat{\theta}$, each based on K failures, were drawn. (The method of drawing the random $\hat{\theta}$'s is explained explicity in the next section.) Then with the n $\hat{\theta}$'s, the first four moments of $g(\theta)$ were estimated as explained in Section 4.3.1, and the K-R method was applied. The criterion for "suitable" data was (roughly) that an (n,K) combination is suitable if the corresponding simulated data leads to a Pearson curve reasonably "close" to the original inverted gamma prior distribution $g(\theta)$. One would expect the estimated prior distribution to become "closer" to the true prior as both n and K increase.

However, the result of the simulation was that, in all 42 cases, the K-R method was not even usable. In all cases, the estimate of the fourth central moment was negative, thus making the K-R method impossible to use (for in all distributions, central moments of even order are positive).

Since the 42 cases are well representative of the practical range of available data, and the specific inverted gamma prior selected is very "typical," the results lead to the following conclusion: No data requirements can be set for the K-R method, since the method generally fails for practical ranges of data.

5.2.2 THE AMOUNT OF DATA - FAMILY SPECIFIED

In this study, two cases of fitting a prior with the family specified were considered:

1) Inverted gamma prior, observed random variable is $\hat{\theta}$. In this case, minimum data requirements must be set for (i) the sample size, n, and (ii) the number of failures $\{K_i\}$, i=1, ..., n.

2) Inverted gamma prior, observed random variable is X, the number of failures in time T. In this case, a minimum data requirement must be set only for the sample size n. (The number of failures in this case is the random variable, not a parameter that can be controlled in a test.)

5.2.2.1 DATA REQUIREMENTS WHEN THE OBSERVED STATISTIC IS $\hat{\theta}$

The general framework is that there is an inverted gamma prior $g(\theta)$on the mean θ of an exponential time-to-failure distribution. Suppose that for each of n equipments, K failure times are observed and a $\hat{\theta}$ is computed. One then fits $g(\theta)$ using the n values of $\hat{\theta}$. (This is the Type 2a data discussed in Section 4.) To test the sensitivity of the fit, one can try fitting the sample data to alternate priors, and can repeat the test for different sets of (n,K). If $g(\theta)$ is the true prior distribution, then intuitively, as n and K are increased, the test should accept $g(\theta)$ as the prior, and reject the alternate distributions, with increasing frequency. In order to get concrete results, i.e., specific values of n and K for which discrimination between the true and alternate priors is "good," it was necessary to conduct simulations for specific cases.

The simulation program reads in the parameters α and λ of $g(\theta)$, along with the parameters of the alternate distributions which were selected to be the Weibull and lognormal distributions with the same mean and variance as the inverted gamma. The program also reads in the desired values of n and K (n = 10,20,30,50,100,200; K = 10,20,30,40). For each of the 24 pairs (n,K), the following steps are carried out:

1) Draw n random observations from $g(\theta)$.

2) For each of the n θ's, draw K random failure times $x_1, \ldots, x_K$ from $f(x|\theta)$ and compute

$$\hat{\theta} = \sum_{i=1}^{K} x_i / K.$$

3) Using the n $\hat{\theta}$'s, perform χ^2 tests to accept or reject each of the 3 candidate priors, at both the .90 and .95 confidence levels (6 χ^2 tests in all).

The whole experiment is repeated 400 times, and the final output is a table that for each n and each K, prints out the number of the 400 that resulted in acceptance, for each of the 6 χ^2 tests.

In order to take the simulated χ^2 tests, it was necessary to compute the percentage points of the marginal distributions corresponding to the inverted gamma, lognormal, and Weibull prior distributions. Since all three marginals are analytically intractable, an auxiliary simulation was written to do the following:

1) 10,000 random samples are drawn from $g(\theta)$.

2) From each of these, K random failure times from an exponential with mean θ are drawn, and a $\hat{\theta}$ is computed.

3) The 10,000 $\hat{\theta}$'s are ordered.

4) The 100th, 200th, ..., 9900th $\hat{\theta}$'s are printed out in a deck to be used in the main program as an approximate "look-up" table of the 99 percentage points of $f_K(\hat{\theta})$.

The χ^2 tests are carried out in the main program as follows:

1) When the n's are read into the program, corresponding C_n's, denoting the desired number of χ^2 cells, and corresponding .90 and .95 level χ^2 values (C_n-1 degrees of freedom), are also read in.

2) For a given (n,K), class marks are selected from the look-up table so that the expected number of observations of $\hat{\theta}$ in each cell is n/C_n under the hypothesis that the marginal distribution $f_K(\hat{\theta})$ is the one corresponding to the inverted gamma prior $g(\theta)$.

3) Using the look-up tables corresponding to the K^{th}-order marginals when the prior is Weibull and lognormal, the expected number of observations in each cell under the Weibull and lognormal hypotheses are found.

4) The number of values of $\hat{\theta}$ falling into each of the cells are tallied and

$$\chi^2 = \sum_{i=1}^{C_n} \frac{(O_i - E_i)^2}{E_i}$$

is computed for all 3 cases, where O_i is the observed number of $\hat{\theta}_K$'s in the i^{th} cell, and E_i the expected number under a given hypothesis. In each case, χ^2 is compared to both the .90 and .95 χ^2 values, and the hypothesis is rejected if χ^2 is too large.

For one of the 400 iterations of the program, a complete output of the χ^2 tests was printed out for each of the 24 (n,K) combinations. Table 5.2.2.1.1 shows the output for the case $n = 100$, $K = 40$. (The inverted gamma parameters are $\alpha = 1500$, $\lambda = 4$). Note that in this particular case the χ^2 test is passed at both levels under the hypotheses that the prior is inverted gamma or lognormal, and is failed at both levels under the Weibull hypothesis.

Due to the fact that the simulation of the percentage points of the marginal distributions $f_K(\theta)$ requires a large amount of computer time, the study was restricted to one inverted gamma prior with parameters $\alpha = 1500$, $\lambda = 4$. The alternate distributions considered were the lognormal with parameters $\mu = 500$, $\sigma = 353.55$, and the Weibull with parameters $\alpha = 8600.7$, $\beta = 1.4355$. All three distributions have the same mean and standard deviation, namely, $\mu = 500$, $\sigma = 353.55$. Plots of the three density functions are shown in Figure 5.2.2.1.2.

The output of the simulation is shown in Tables 5.2.2.1.3.a and 5.2.2.1.3.b. For each pair (n,K), 400 simulations were run, taking 2 χ^2 tests (.90 and .95 level) for each of the above 3 priors. The number of tests passed out of 400 are given for the .90 level tests in Table 5.2.2.1.3.a and for the .95 level tests in Table 5.2.2.1.3.b.

Referring to the above tables, the following rationale for selecting suitable (n,K) pairs is given:

NOT REPRODUCIBLE

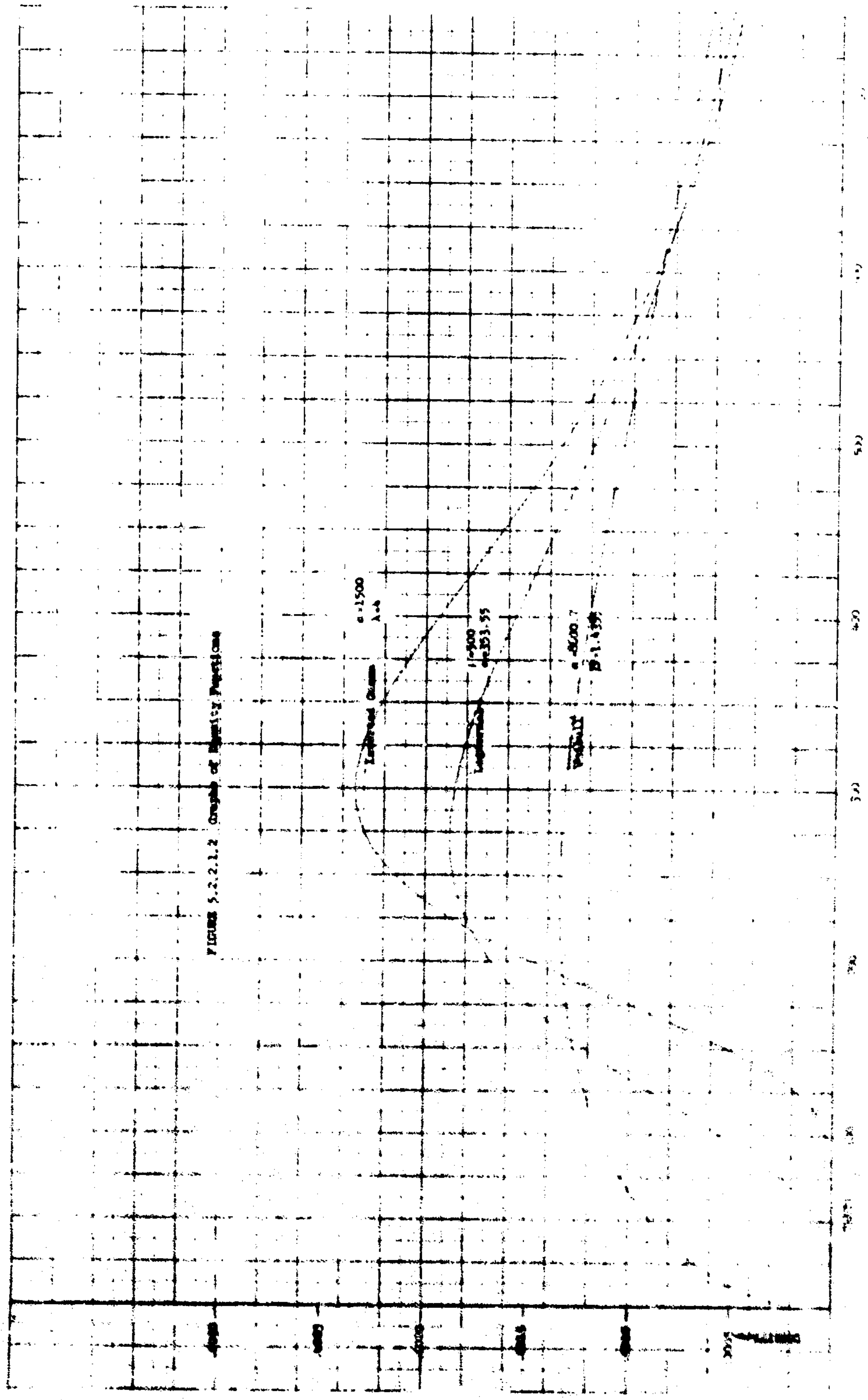


FIGURE 5.2.2.1.2 Graphs of Density Functions

TABLE 5.2.2.1.1 EXPECTED AND OBSERVED VALUES FOR χ^2 TESTS UNDER EACH OF 3 HYPOTHESIZED PRIORS

$n = 100$

$K = 40$

Cell No.	Upper Class Mark for Cell	Observed No. of $\hat{\theta}$'s	Expected No. for Inverted Gamma Prior	Expected No. for Log Normal Prior	Expected No. for Weibull Prior
1	215.91	15	10	17.634	23.686
2	260.69	4	10	8.488	6.150
3	304.44	10	10	7.791	6.169
4	354.52	16	10	8.163	5.985
5	404.79	9	10	7.691	6.147
6	469.57	10	10	9.145	7.444
7	553.02	12	10	8.804	8.065
8	669.81	7	10	9.425	9.346
9	884.06	7	10	10.395	12.427
10		10	10	12.465	14.581

χ^2 VALUES

	Inverted Gamma	Log Normal	Weibull
Computed χ^2	12.000	14.598	31.592
.90 level	14.684	14.684	14.684
.95 level	16.919	16.919	16.919

If N χ^2 tests are run at significance level p, then the distribution of the number of tests passed, under the null hypothesis H_0 that the samples come from the distribution being fitted, is binomial with mean $\mu = Np$ and variance $\sigma^2 = Np(1-p)$.

In case a, $N = 400$, $p = .90$, so that $\mu = 360$, $\sigma^2 = 36$.

In case b, $N = 400$, $p = .95$, so that $\mu = 380$, $\sigma^2 = 19$.

In each case, letting the random variable X be the number of tests passed, the distribution of $(X-\mu)/\sigma$ is approximately distributed N(0,1)(unit normal). If we solve the equations

$$\frac{x_{.90}-\mu}{\sigma} = -1.28 \quad \text{and} \quad \frac{x_{.95}-\mu}{\sigma} = -1.64$$

for $x_{.90}$ and $x_{.95}$, then

$$\Pr(x \geq x_{.90}|H_0) = .90 \text{ and } \Pr(x \geq x_{.95}|H_0) = .95.$$

Referring to Tables 5.2.2.1.3.a and 5.2.2.1.3.b, the following criterion seems reasonable: "Suitable" pairs (n,K) for fitting the prior are those for which the number of "passes" exceed $x_{.90}$ (or $x_{.95}$) for the test of the inverted gamma prior (the true prior) and are less than $x_{.90}$ (or $x_{.95}$) for the tests of the lognormal and Weibull priors (the false hypotheses). The computed values of $x_{.90}$ and $x_{.95}$ (rounded to the nearest integer) are:

	$x_{.90}$	$x_{.95}$
Case a (.90 level)	352	350
Case b (.95 level)	374	373

Looking at Table 5.2.2.1.3.a, one can note the following:

1) Excluding the case where n = 10 (all K), and the two cases n = 20, K = 5 and n = 20, K = 10, the number of passes for the lognormal and Weibull tests is always less than both $x_{.90}$ and $x_{.95}$ (352 and 350), whereas, for the inverted gamma test, only 4 cases pass less than 352 times, and only 1 less than 350 times.

2) For each fixed K, the number of times the lognormal and Weibull tests pass monotonically decreases as n increases (with one exception). The same holds true (with 2 exceptions) for fixed n $\neq$ 10 and increasing K.

TABLE 5.2.2.1.3.a. NUMBER OF χ^2 TESTS PASSED (OUT OF 400) UNDER EACH OF 3 HYPOTHESIZED PRIORS, TAKEN AT .90 LEVEL

K \ n		10	20	30	50	100	200
5	Inverted Gamma	357	375	364	351	367	360
	Lognormal	357	359	342	311	308	268
	Weibull	357	320	272	175	84	14
10	Inverted Gamma	364	365	364	345	357	351
	Lognormal	364	351	347	302	251	210
	Weibull	364	299	233	131	31	0
20	Inverted Gamma	356	359	365	358	359	359
	Lognormal	356	330	319	313	239	158
	Weibull	356	251	172	86	9	0
40	Inverted Gamma	356	367	359	365	367	349
	Lognormal	356	318	298	261	189	85
	Weibull	356	244	161	75	7	0

An investigation of Table 5.2.2.1.3.b yields essentially the same results as above.

As can be seen in the tables, the discrimination between the true and alternate priors increases rather dramatically as n and K increase. It can be seen why the Weibull distribution fares "worse" than the lognormal by looking at the plot of the three densities (Figure 5.2.2.1.2). Obviously, the less the alternate prior "looks like" the true prior, the better the discrimination between the two.

On the basis of the above results, the following conservative data requirements are set for fitting an inverted gamma prior when the observed statistic is $\hat{\theta}$: $n > 20$ and $K > 10$. We must keep in mind, of course, that the requirements are based on experience with one particular prior (inverted gamma with $\alpha = 1500$, $\lambda = 4$) and two particular alternate priors. It would be of interest to repeat the simulation for other cases. It is felt, however, that the case considered is fairly representative and that the results obtained from it are generally applicable.

The only "problem" with the analysis described above is that it is limited to cases where K is fixed in each sample of $\hat{\theta}$'s (Type 2a data). When the simulations were first undertaken, it was thought that if a pair (n,K) was suitable for a prior fit, then data of the form $\{(\hat{\theta}_i, K_i), i=1, \ldots, n$, all $K_i \geq K\}$ would also be suitable. This is not the case, however, since when the K_i's are different (Type 2b data), one must use the mixed population model for the marginal distribution. As explained in Section 4.3.2, this leads, in actual practice, to a problem concerning the degrees of freedom for the χ^2 test, due to n being too small and/or having too much "variety" in the K_i's. In the above simulation, this problem was avoided by having K's for each sample. The basic quantitative data requirement when the K_i's are different is that derived in Section 4.3.2: $c \geq t + 4$, where c is the number of χ^2 cells, and t is the number of distinct K_i's. Since we require $\frac{n}{5} \geq c$, we can rewrite the requirement as $t + 4 \leq \frac{n}{5}$.

This requirement was not achieved in this report as can be seen in Section 4.3.2. Two remarks are in order. First, no matter how many distinct K_i's there are (even if t=n) if the K_i are all large, i.e., greater than or equal to 20, then the prior distribution can be fitted directly to the observed $\hat{\theta}$'s. Secondly, when n is relatively large with respect to t the mixed model can be fitted. There will be cases for large n when t is relatively small.

5.2.2.2 DATA REQUIREMENTS WHEN THE OBSERVED STATISTIC IS THE NUMBER OF FAILURES IN TIME T

Sections 4.2.2.1 and 4.2.3.2 discuss the method, and Section 4.3.2 the results, of fitting an inverted gamma prior distribution when the observed statistic is the number of failures X in a fixed time T. The purpose of

TABLE 5.2.2.1.3.b. NUMBER OF χ^2 TESTS PASSED (OUT OF 400) UNDER EACH OF 3 HYPOTHESIZED PRIORS, TAKEN AT .95 LEVEL.

K \ n		10	20	30	50	100	200
5	Inverted Gamma	394	390	382	379	381	376
	Lognormal	394	377	373	343	348	299
	Weibull	394	349	318	233	124	22
10	Inverted Gamma	393	385	381	381	376	377
	Lognormal	381	373	372	343	306	259
	Weibull	393	335	286	186	53	1
20	Inverted Gamma	392	379	380	378	382	379
	Lognormal	377	364	357	343	290	203
	Weibull	392	286	229	127	17	0
40	Inverted Gamma	388	380	376	383	387	364
	Lognormal	388	360	338	320	250	130
	Weibull	376	284	217	114	17	0

this section is to find the minimum unit sample size n required to get a "suitable" prior fit. The analysis to determine the minimum n was based on the sensitivity study described below, which is analogous to the one described in the previous section. The restriction on T is that it be large enough so that the number of failures is variable enough to take the χ^2 test.

The inverted gamma prior parameters selected for the simulation were $\alpha = 4000$, $\lambda = 3$. As in Section 5.2.2.1, two alternate priors were selected having the same mean and variance as the above prior: a Weibull distribution with parameters $\alpha = 2000$, $\beta = 1$ and a lognormal distribution with parameters $\mu = 2000$, $\sigma = 2000$. The fixed time T was chosen to be $T = 4000$ hours. Six values for the sample size n were studied: $n = 10,20,30,50,100,200$. For each n, n random values of X, $\{x_1, \ldots, x_n\}$ were drawn, and χ^2 tests taken against the (true) inverted gamma prior and the alternate hypothesized priors. Each random value of X was drawn by taking a random value of θ from the inverted gamma prior $g(\theta)$, and then taking random failure times $t_1, t_2, \ldots$ from the exponential distribution with mean θ. The value of X is then taken to be the largest value of K for which $\sum_{i=1}^{K} t_i < T$.

Table 5.2.2.2.1 shows the results of one run of the simulation for the case $n = 200$. The first column of the upper table lists the value, x, that the random variable X can take on, and the second column gives the observed number of times (out of 200) that the events $X=x$ occurred. The last 3 columns give the expected number of times (out of 200) of the occurrence of the events $X=x$ under the three hypothesized prior distributions. These expectations are easy to compute for the case of the inverted gamma prior, since the marginal distribution of X can be computed explicitly by Formula 4.2.2.1.2. For the lognormal and Weibull cases, however, the marginal distributions had to be approximated by simulation. In both cases, 10,000 random values of X were drawn (by taking random θ's from $g(\theta)$ and obtaining a random value of X as described in the previous paragraph). Then the relative frequencies (out of 10,000) of the events $X = 0,1,2,\ldots,$ were used to approximate the marginal $f(x)$.

In the lower table of Table 5.2.2.2.1, the results of the χ^2 tests are given. Six degrees of freedom were used because the data in the upper table was divided into 7 χ^2 cells. For this particular case, the inverted gamma prior passes the χ^2 test at both the .90 and .95 level, while the other two hypothesized priors fail both tests.

As in Section 5.2.2.1, the above simulation was repeated 400 times for each n. Table 5.2.2.2.2 gives the number of times out of 400 that each test was passed. The criterion for "suitable" values of n is the same as in Section 5.2.2.1, leading to the same critical values $x_{.90}$ and $x_{.95}$ as given in that section. As can be seen in the table, when n is ≥ 30, the number of times the inverted gamma prior passes the χ^2 test exceeds the critical value 350(373) at the .90(.95) level, whereas, the number of passes for the lognormal and Weibull priors fall short of the critical value.

Hence, in the case of fitting a prior specified as inverted gamma when the observed data is the number of failures in time T, the following conservative data requirement is set: the sample size n must be at least 30.

TABLE 5.2.2.2.1 EXPECTED AND OBSERVED VALUES FOR χ^2 TESTS UNDER EACH OF 3 HYPOTHESIZED PRIORS (n=200)

n = 200

Value x of X	Observed No. of Times X = x	Expected No. for Inverted Gamma Prior	Expected No. for Log Normal Prior	Expected No. for Weibull Prior
0	33	25.00	25.16	28.76
1	43	37.50	35.06	34.20
2	41	37.50	32.08	26.80
3	29	31.25	24.92	20.68
4	13	23.44	19.60	14.52
5	16	16.41	15.80	10.20
6	8	10.94	11.34	7.70
7	11	7.03	8.24	7.12
Over 7	6	10.94	22.06	44.62

χ^2 VALUES

	Inverted Gamma	Lognormal	Weibull
Computed χ^2	8.964	44.196	20.034
.90 Level	10.645	10.645	10.645
.95 Level	12.592	12.592	12.592

TABLE 5.2.2.2.2 NUMBER OF χ^2 TESTS (OUT OF 400) UNDER EACH OF 3 HYPOTHESIZED PRIORS

χ^2 Level	n	10	20	30	50	100	200
.90 Level	Inverted Gamma	352	342	353	355	353	367
	Lognormal	352	332	328	260	183	57
	Weibull	373	332	263	46	11	0
.95 Level	Inverted Gamma	390	380	383	373	376	387
	Lognormal	373	364	352	307	228	89
	Weibull	373	364	304	80	16	0

SECTION 6.0 ANALYSIS FOR DATA COMBINATION

In fitting prior distributions in this study different classes of identical equipments were available. In fact, seventeen data sets were used. It turns out that some care must be exercised in defining what is meant by a prior distribution on MTBF, say $g(\theta)$. All p.d.f.'s must exhibit some form of homogeneity. That is, limitations are placed on the "reasons" why the variable of interest varies. These reasons are commonly called assignable causes of variation. As a beginning, $g(\theta)$ has been restricted as follows:

> The random MTBF's (θ's) belonging to a particular $g(\theta)$ must be θ's on a given type equipment built to the same design specifications by a particular manufacturer.

This is somewhat restrictive because it apparently places equipment of the same design but built by different manufacturers in different $g(\theta)$'s. It also places similar equipment, say computers, but with different designs, e.g., different memory size, in different $g(\theta)$'s. One might ask, couldn't some of this data, say different computer types, be combined into one prior distribution? The answer is yes, but with this qualification: every assignable cause of variation, e.g., different manufacturers, is a piece of prior information and, if possible, should be exploited. Combining data into one prior which have assignable and identifiable causes for having different MTBF's, in general, increases the variation in the prior distribution. On the other hand, fitting a large number of prior distributions is a costly process and it is worthwhile to be able to "relate" prior distributions on similar equipments even though they, the prior distributions, are not combined. For example, two computers of similar design, say different memory sizes, might have prior distributions which are relatable, though it might not be wise to combine them. Thus, when the prior distribution of the one is fitted, the prior distribution of the other is known. For example, let these computers be called c_1, c_2 with prior distributions $g_1(\theta)$ and $g_2(\theta)$. It might turn out that $g_1(\theta)$ and $g_2(\theta)$ belong to the same family, e.g., inverted gamma, and hence may be related by some transformation. If this transformation is known then having fitted $g_1(\theta)(g_2(\theta)), g_2(\theta)(g_1(\theta))$ is known. In Section 9.3 of the Appendix, a particularly simple transformation was considered, namely,

$$\theta_2 = K\theta_1 \qquad K > 0.$$

Then, if θ_1 is inverted gamma with parameters (α_1, λ_1), θ_2 is inverted gamma with parameters $(\alpha_2 = \alpha_1 K, \lambda_2 = \lambda_1)$. Thus, knowing (α_1, λ_1) and the constant K, the parameters of the second prior distribution are known.

One particularly good idea for estimating K, which could not be verified because of lack of data on two similar equipments, is that K might be well represented by

$$K = \theta_{p_2}/\theta_{p_1}$$

where θ_{p_i} is the predicted MTBF on the i^{th} equipment. This idea should be investigated in the next study phase.

SECTION 7.0 ROBUSTNESS ANALYSIS

SUMMARY

Objective

The objective of this section was to investigate the effects of errors in estimating the scale and shape parameters of inverted gamma prior distribution on the posterior inverted gamma distribution. The effects were measured in terms of changes in the mean, 5^{th}, 10^{th}, 90 and 95^{th} percentiles of the posterior inverted gamma distribution. For measuring effects of errors in estimating the scale parameter α, a shape parameter $\lambda=3$ was assumed. For measuring effects of errors in estimating the shape parameter a scale parameter $\alpha=200$ was assumed. The posterior distribution depends on $\hat{\theta}$ and three values were chosen: $\hat{\theta}=50$, 100, 200. These choices were reasonable in terms of the value of α,λ assumed; K values chosen were 5,10,20,30.

Results

The effect of errors in estimating the scale parameter of the prior inverted gamma distribution were practically negligible on the selected percentiles of the posterior inverted gamma distribution for all three values of $\hat{\theta}$ and for $K \geq 20$. The effects were more noticeable for K=10 and quite pronounced for K=5. This ties in with the results of Section 5.0 regarding data requirements although arrived at in a different way.

The effects of errors in estimating the shape parameter λ in the prior inverted gamma distribution were, as in the above case, practically negligible on the selected percentiles for $K \geq 20$. For K = 5,10 the effects were quite pronounced. This result again ties in with the results of Section 5.0.

It had originally been intended to select several Bayes reliability demonstration test methods and investigate their robustness with respect to errors in estimating the parameters and family of the prior distribution. However, only two methods (Bibliography #2, #59) are available at this time and neither has yet gained anything near acceptance in the field of reliability testing. For these reasons, it was decided to abandon this approach and take another tack which, hopefully, would be of more use to the reader.

The ultimate use of the prior distribution is to supplement it with observed data so that a posterior distribution may be calculated. In short, Bayes reliability tests will use the posterior distribution in one form or another. Thus, it was decided to investigate the sensitivity of posterior distribution to errors in estimating the parameters of the prior inverted gamma distribution; there was not time to investigate differences in families.

If (α, λ) are the scale and shape parameters of a prior inverted gamma distribution, then the posterior distribution has parameters $(\alpha + K\hat{\theta}, \lambda + K)$ with mean

$$E(\theta|\hat{\theta}) = \frac{\alpha+K\hat{\theta}}{\lambda+K-1} \tag{7.1}$$

Now, if (α,λ) were estimated incorrectly as say, $\alpha' = \alpha + \Delta\alpha$, $\lambda' = \lambda + \Delta\lambda$ then

$$E'(\theta|\hat{\theta}) = \frac{\alpha' + K\hat{\theta}}{\lambda' + K-1} = \frac{\alpha + \Delta\alpha + K\hat{\theta}}{\lambda + \Delta\lambda + K-1} \tag{7.2}$$

and this would be the posterior mean used. However, for large K both $E(\theta|\hat{\theta})$ and $E'(\theta|\hat{\theta})$ approach the same limit, i.e., $\hat{\theta}$, and do not differ by much. The percentiles of the posterior distribution are not available in closed form but can easily be obtained by computer. Figures 7.1 through 7.12 present the behavior of the posterior inverted gamma distribution (in terms of its mean, 5th, 10th, 90th and 95th percentiles) for a fixed value of $\lambda = 3$ as a function of the scale parameter α. The $\lambda = 3$ was selected because it represented a reasonable magnitude for the shape parameter. Since the posterior distribution depends on $\hat{\theta}$, three $\hat{\theta}$'s were selected for each K: 50, 100, 200. The K's selected were 5, 10, 20, 30. Figure 7.1 shows that the mean and percentiles of the posterior distribution are rather insensitive to α (i.e., the lines are almost vertical) for $K = 30$, $\hat{\theta} = 200$. Much the same is true for all $K = 30$, i.e., for Figures 7.1, 7.2, 7.3. When $K = 10$ (Figures 7.7, 7.8, 7.9) the mean and percentiles become quite sensitive to differences in α. Note that the effects of $\hat{\theta}$ are in the translation of the mean and percentiles along the θ axis and not in the slope of the lines. Thus, for $\lambda = 3$ the errors of estimating α incorrectly do not appear serious until K gets small.

Comparing Figures 7.1 through 7.12 with Figure 7.0 shows that though the prior distribution is highly sensitive to differences in α (for $\lambda = 3$) the posterior distributions do not exhibit this characteristic. The posterior distributions are, of course, sensitive to differences in $\alpha + K\hat{\theta}$ but $K\hat{\theta}$ is always known and usually large compared to α so that errors in estimating α are not too serious.

A study similar to the previous discussion was done for the same $\hat{\theta} = 50, 100, 200$, $K = 5, 10, 20, 30$ but this time $\alpha = 200$ was held fixed and λ varied. Figure 7.13 shows the prior distribution mean and percentiles. Since the shape parameter of the posterior distribution (Figures 7.14 through 7.25) is $(\lambda + K)$, it is no surprise that the sensitivity of the mean and percentiles to errors in λ is greater for smaller K. For large K, $(\lambda + K)$ is dominated by K for the values of λ expected to occur in reliability. Put another way, suppose $\lambda = 5$ in the prior distribution but that an 60% error was made and $\lambda' = 8$ was estimated. For $K = 30$ then the true shape parameter in the posterior distribution would be $(5 + 30) = 35$ and the one estimated would be $(8 + 30) = 38$ a relatively small percentage error compared to the original error.

Figure 7.[illegible] Mean and Selected Percentiles-Prior Inverted Gamma Distribution

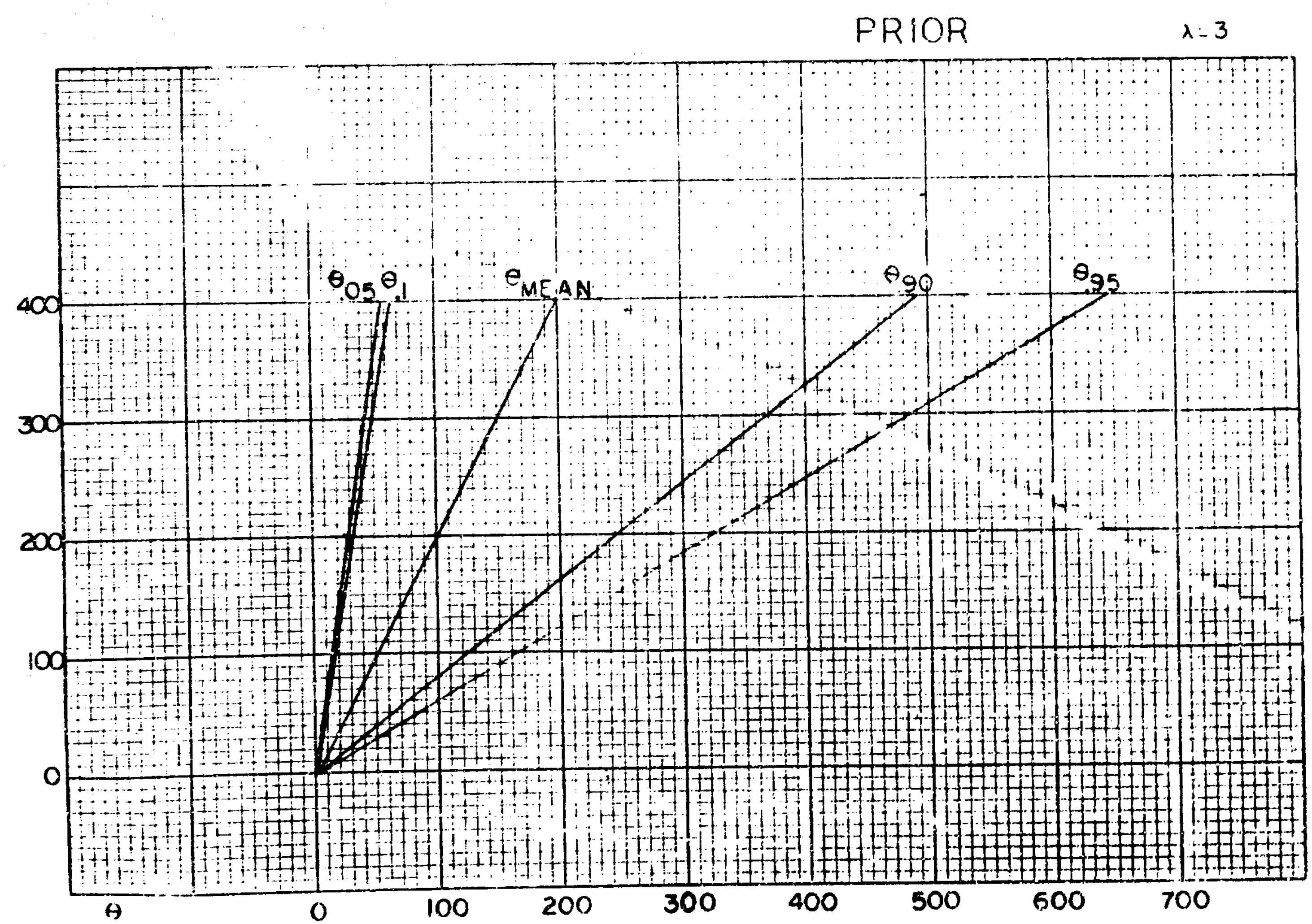


Figure 7.1 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

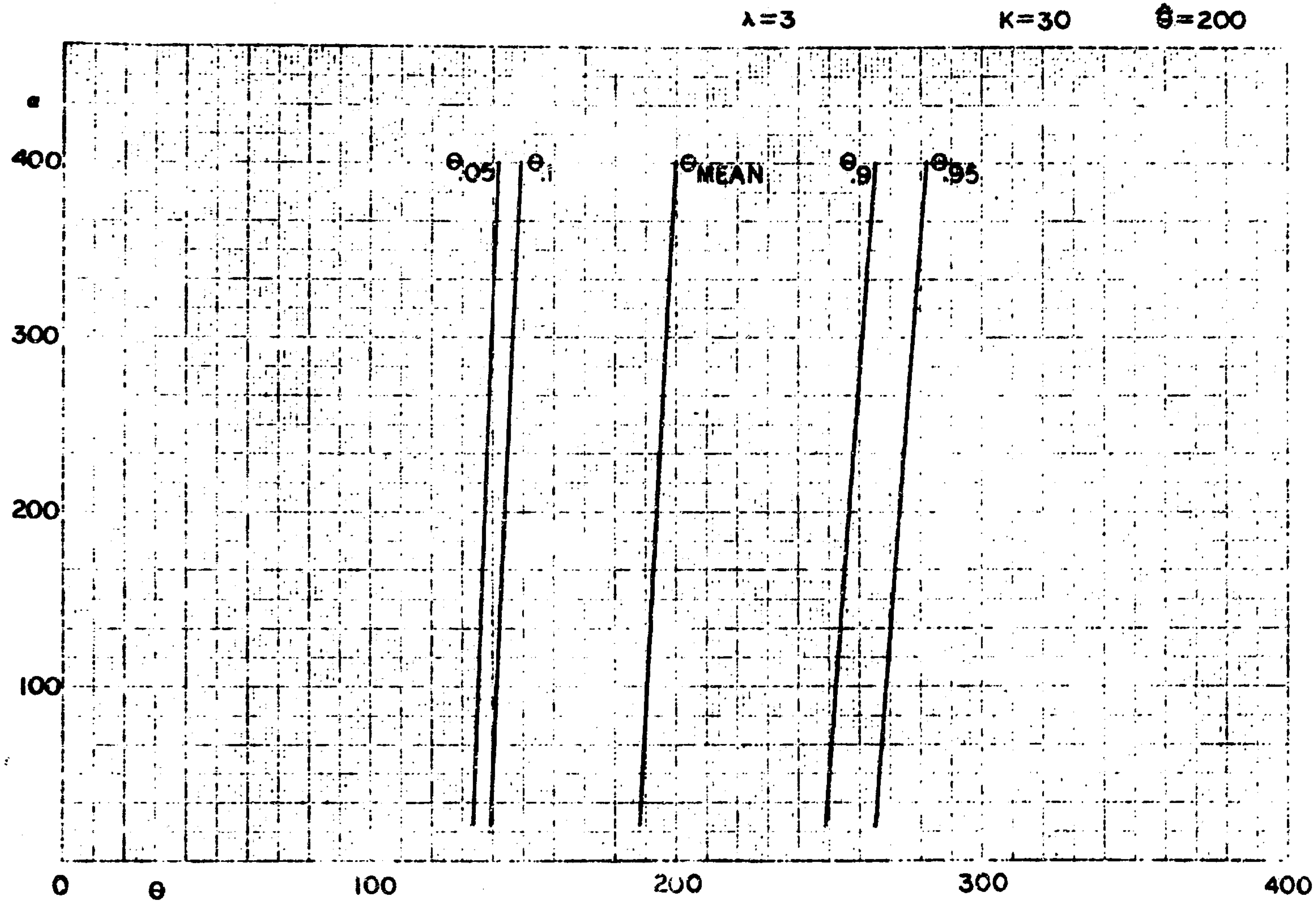


Figure 7.2 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda=3$ $K=30$ $\hat{\theta}=100$

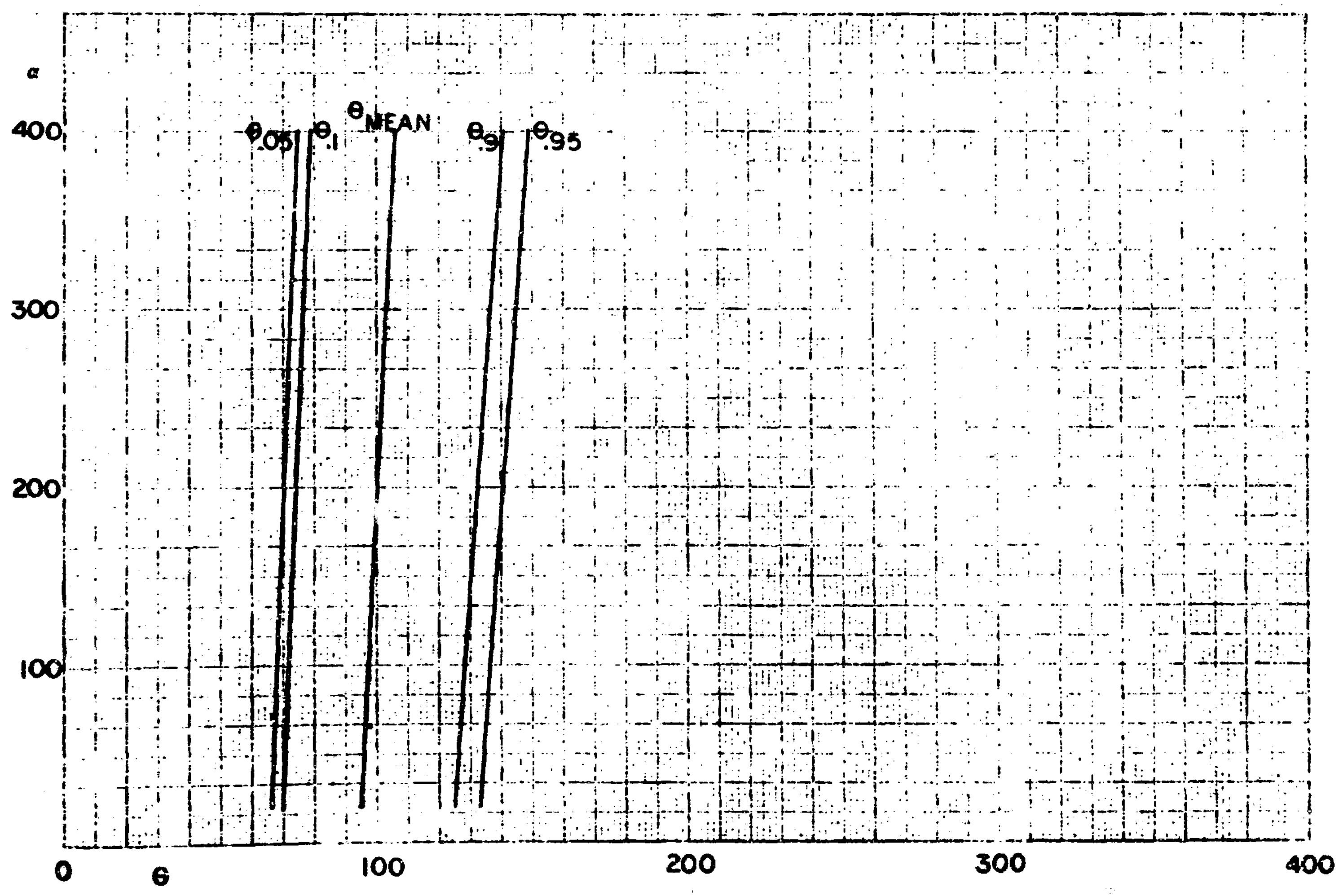


Figure 7.3 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda=3$ $K=30$ $\hat{\theta}=50$

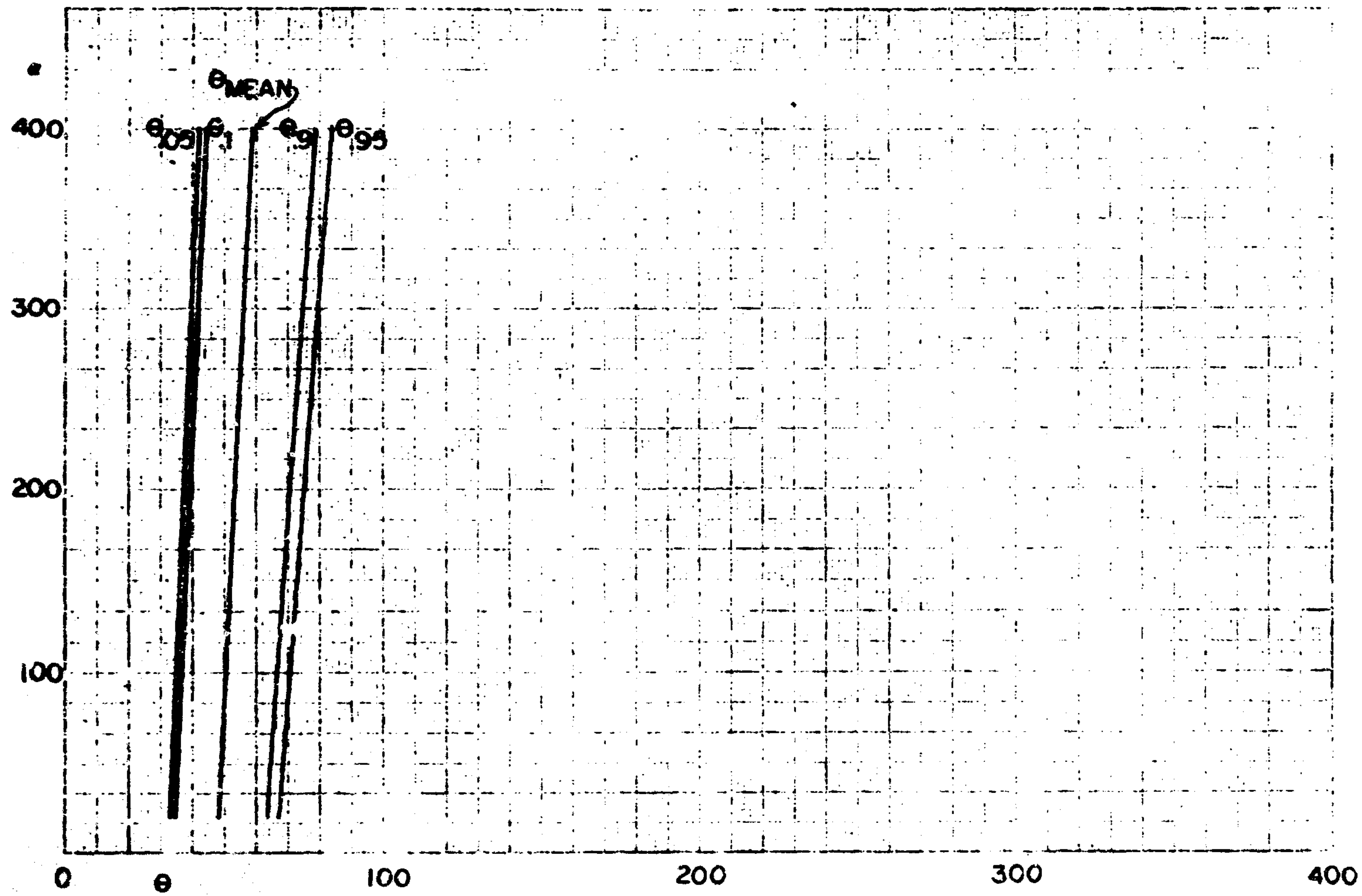


Figure 7.4 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda=3$ $K=20$ $\hat{\theta}=200$

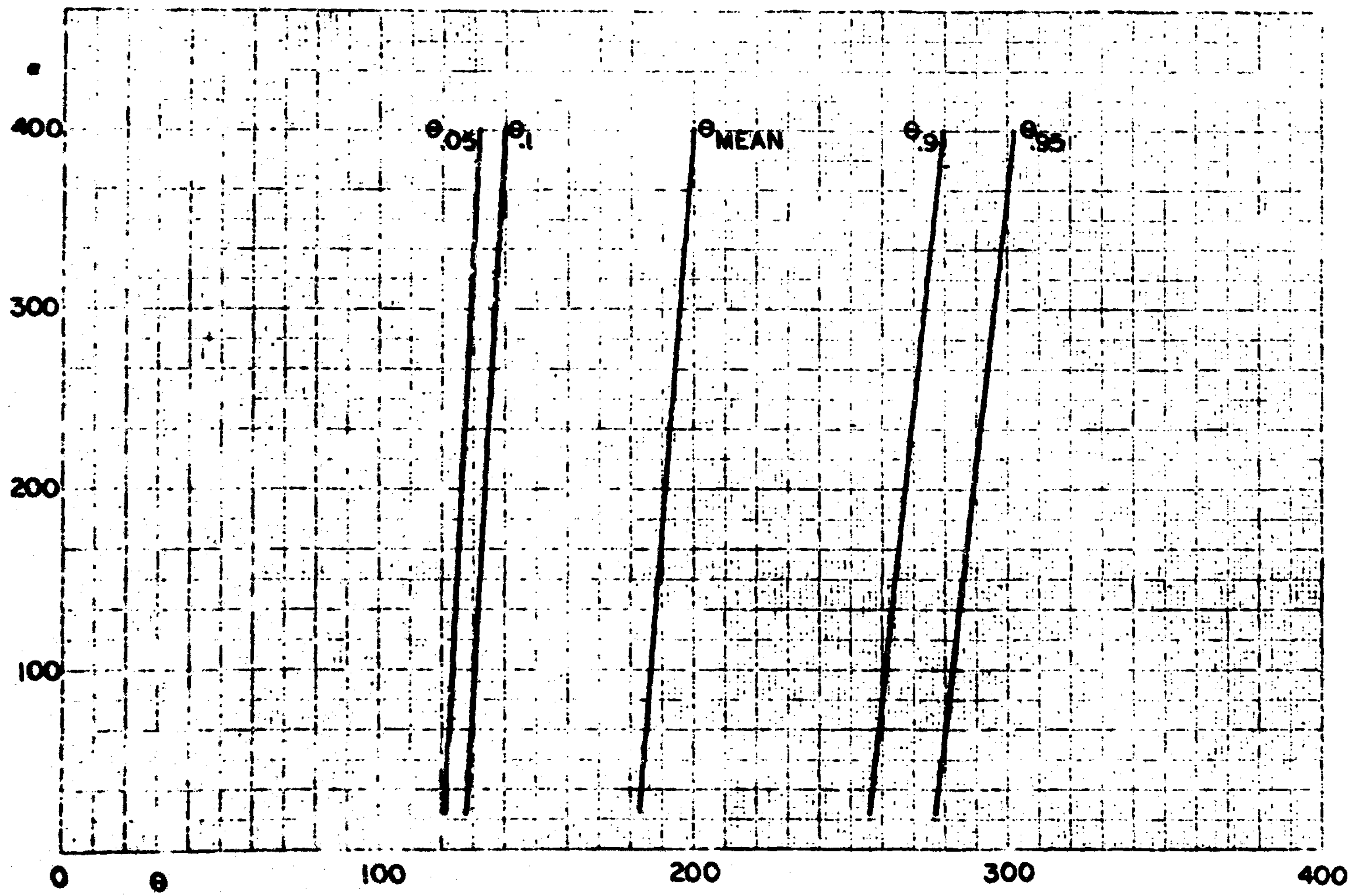


Figure 7.5 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

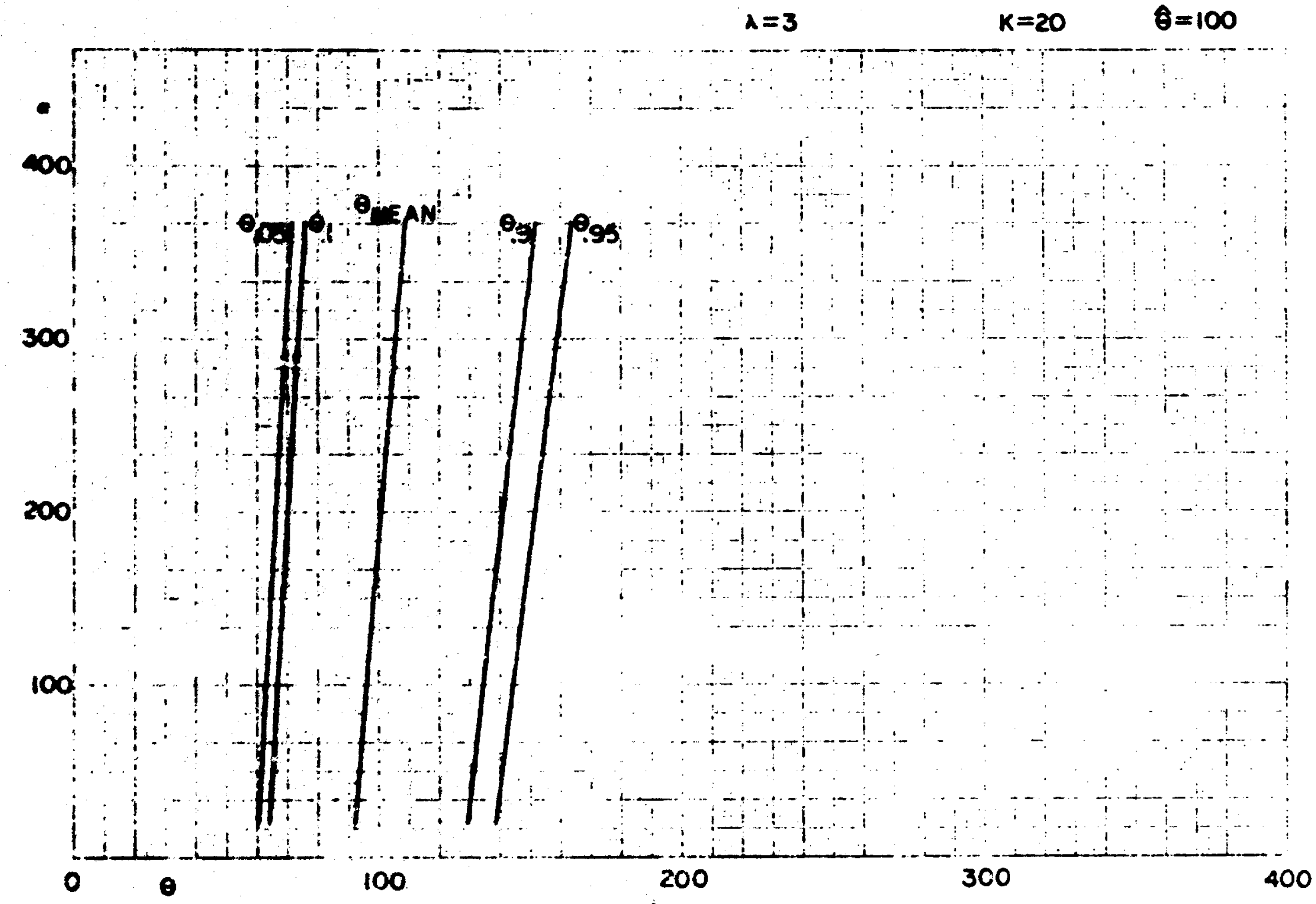


Figure 7.6 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

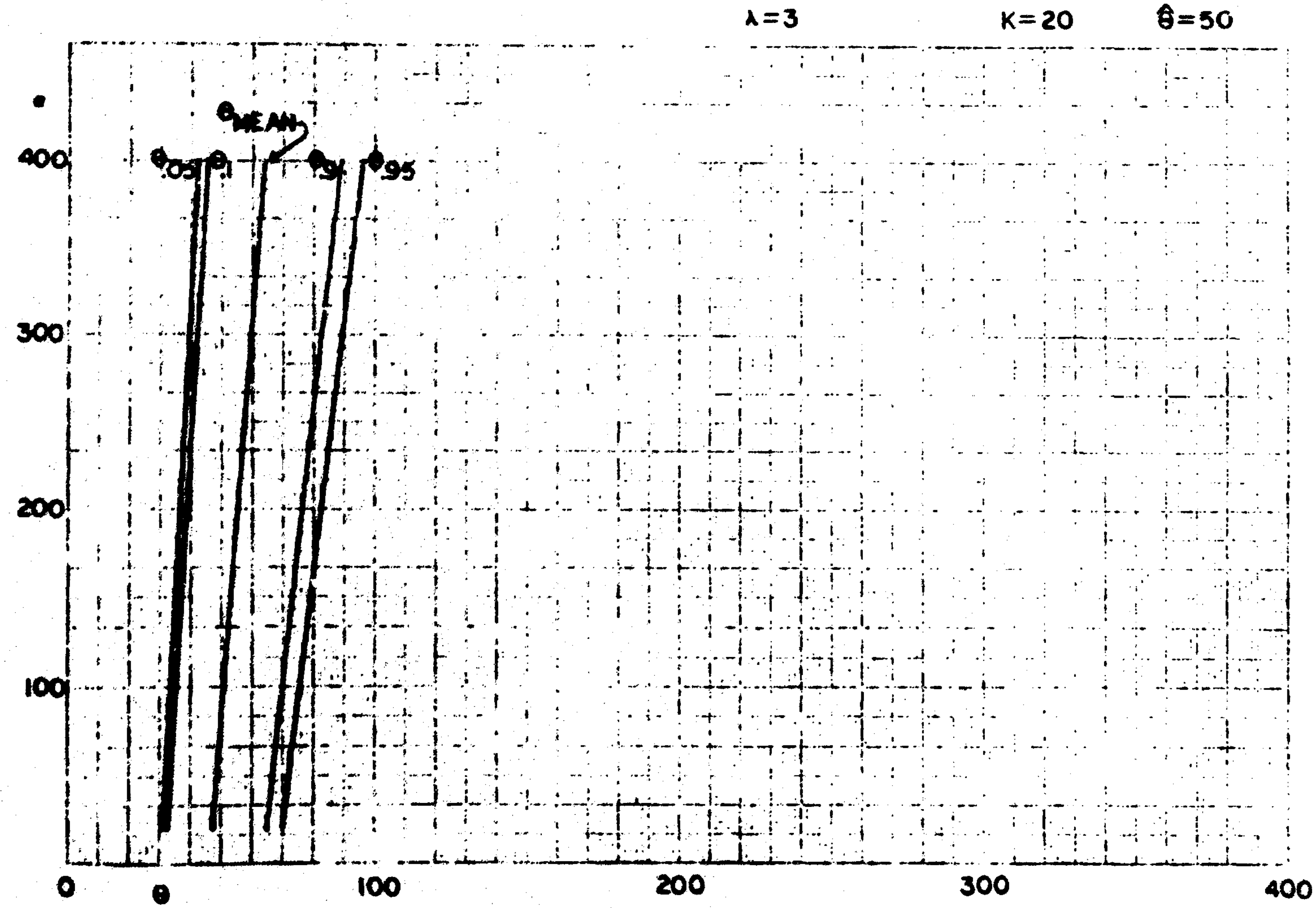


Figure 7.7 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

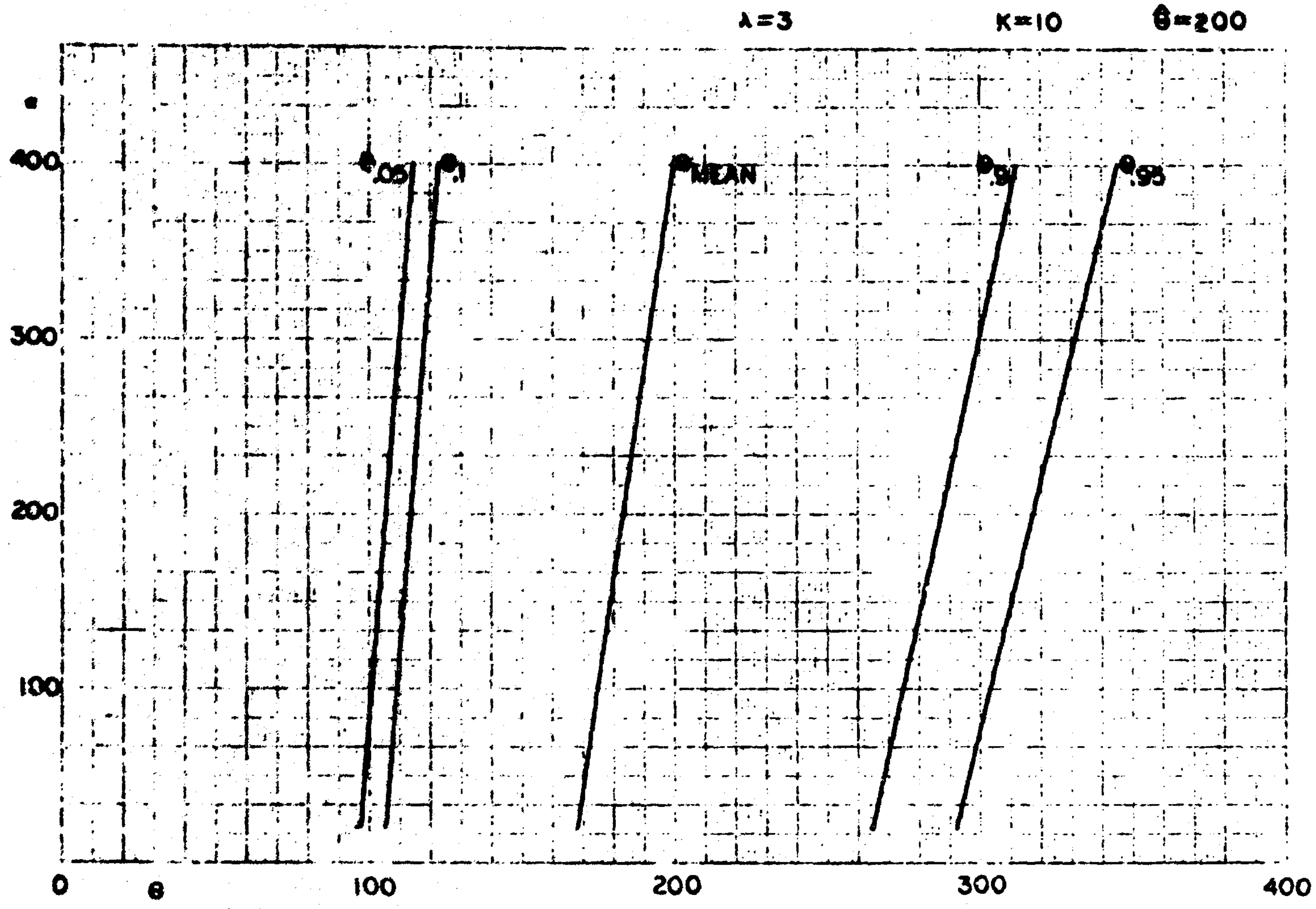


Figure 7.8 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda = 3$ $K = 10$ $\hat{\theta} = 100$

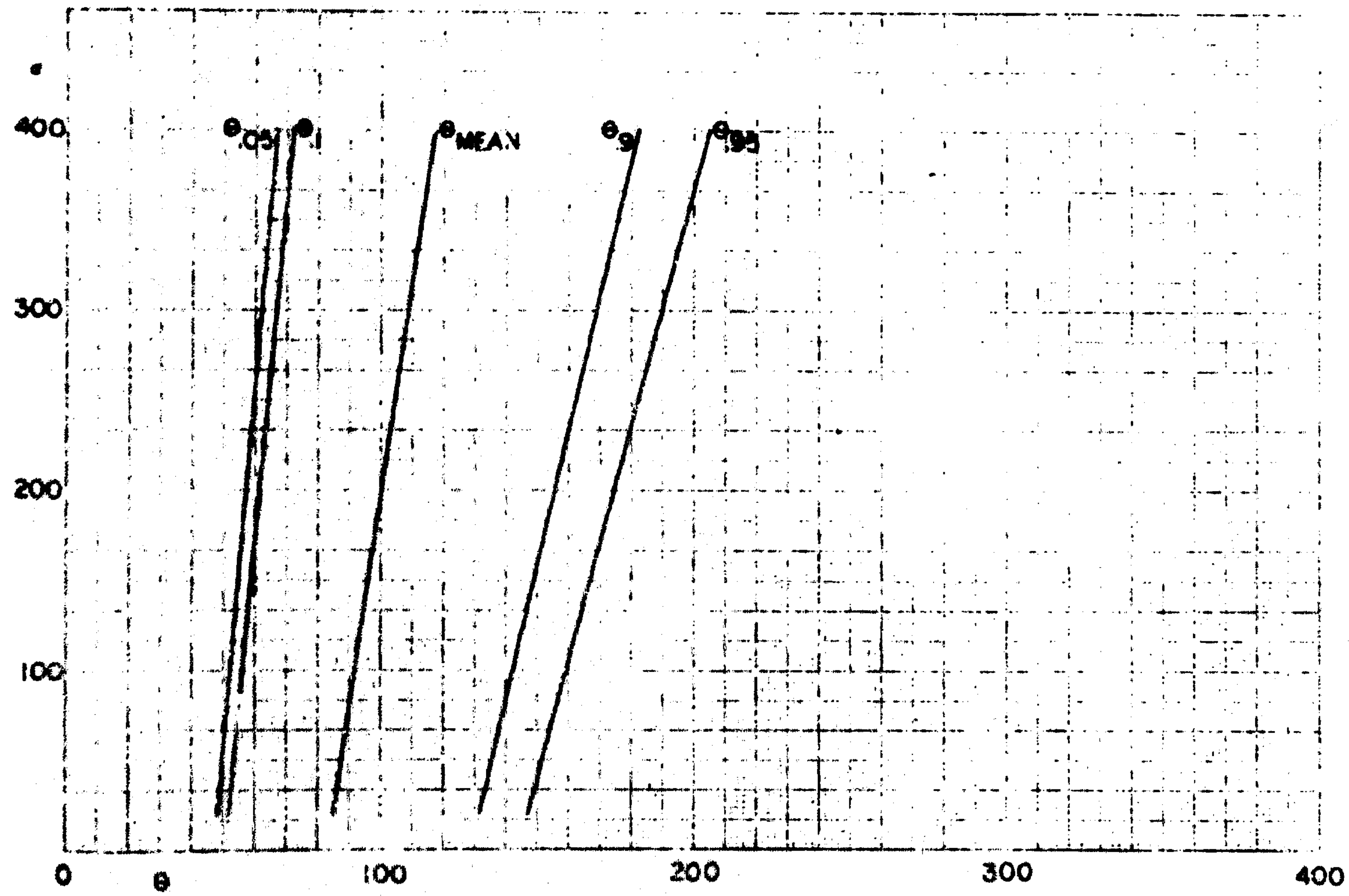


Figure 7.9 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

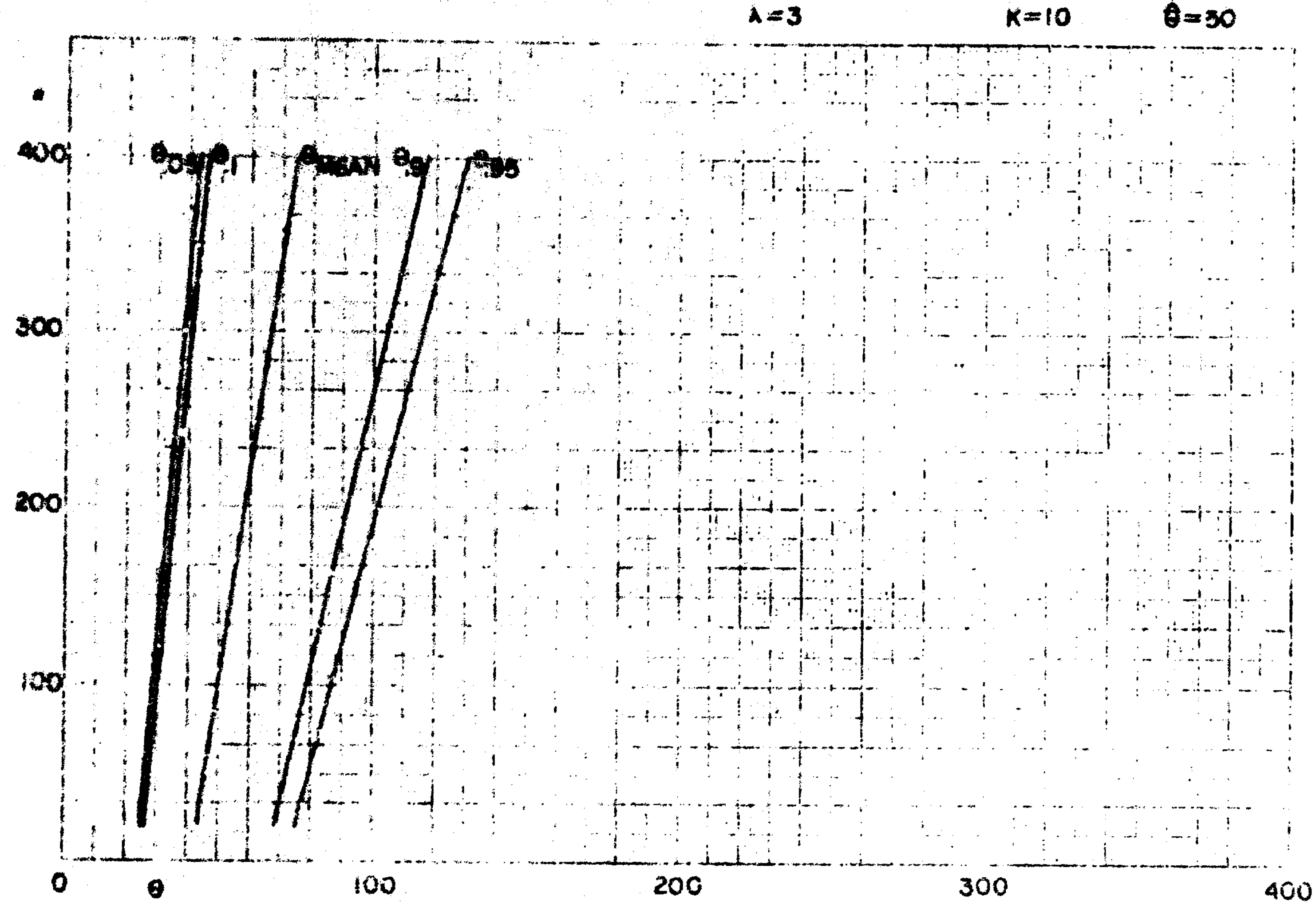


Figure 7.10 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda = 3$ $K = 5$ $\hat{\theta} = 200$

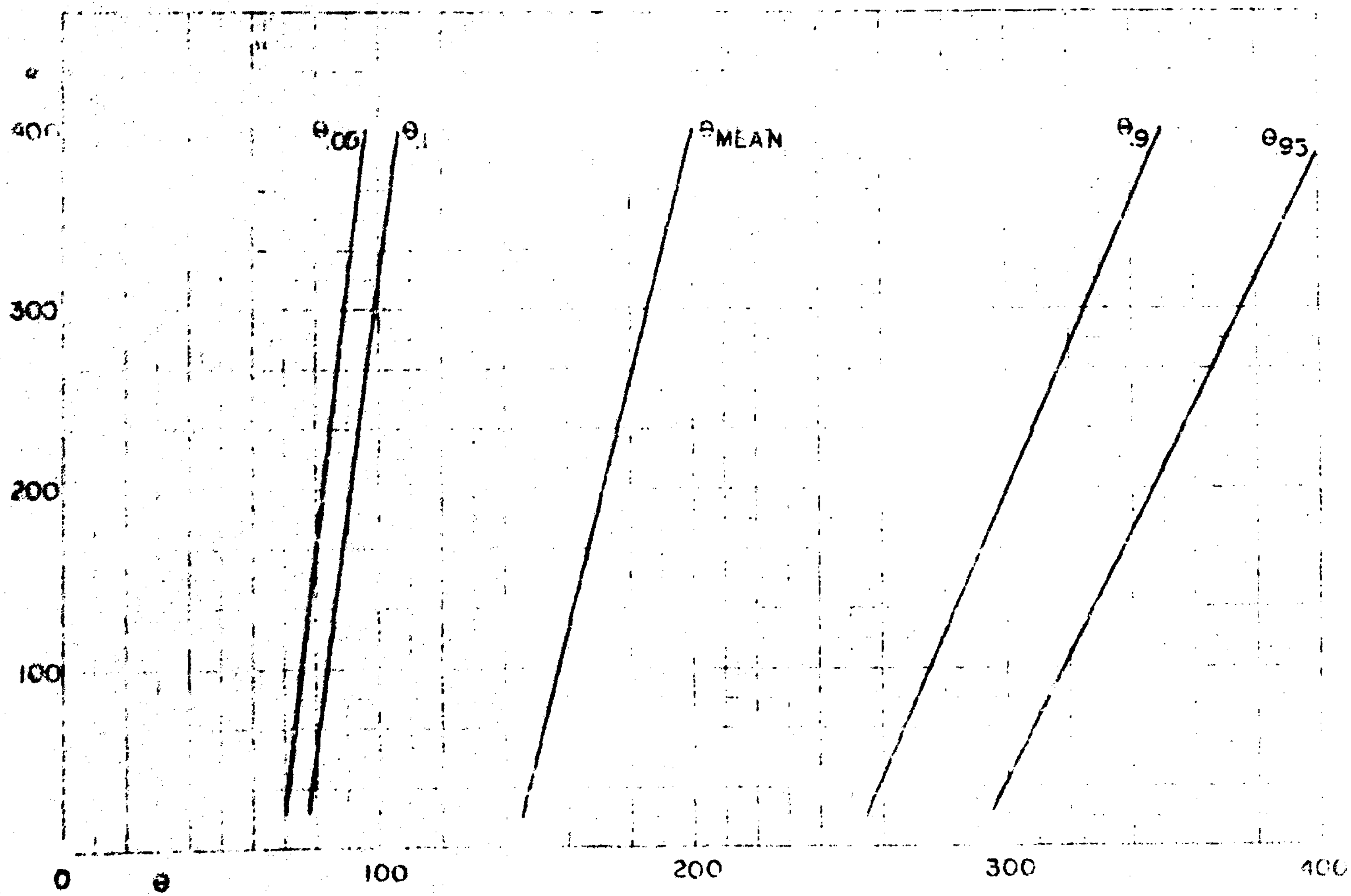


Figure 7.11 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

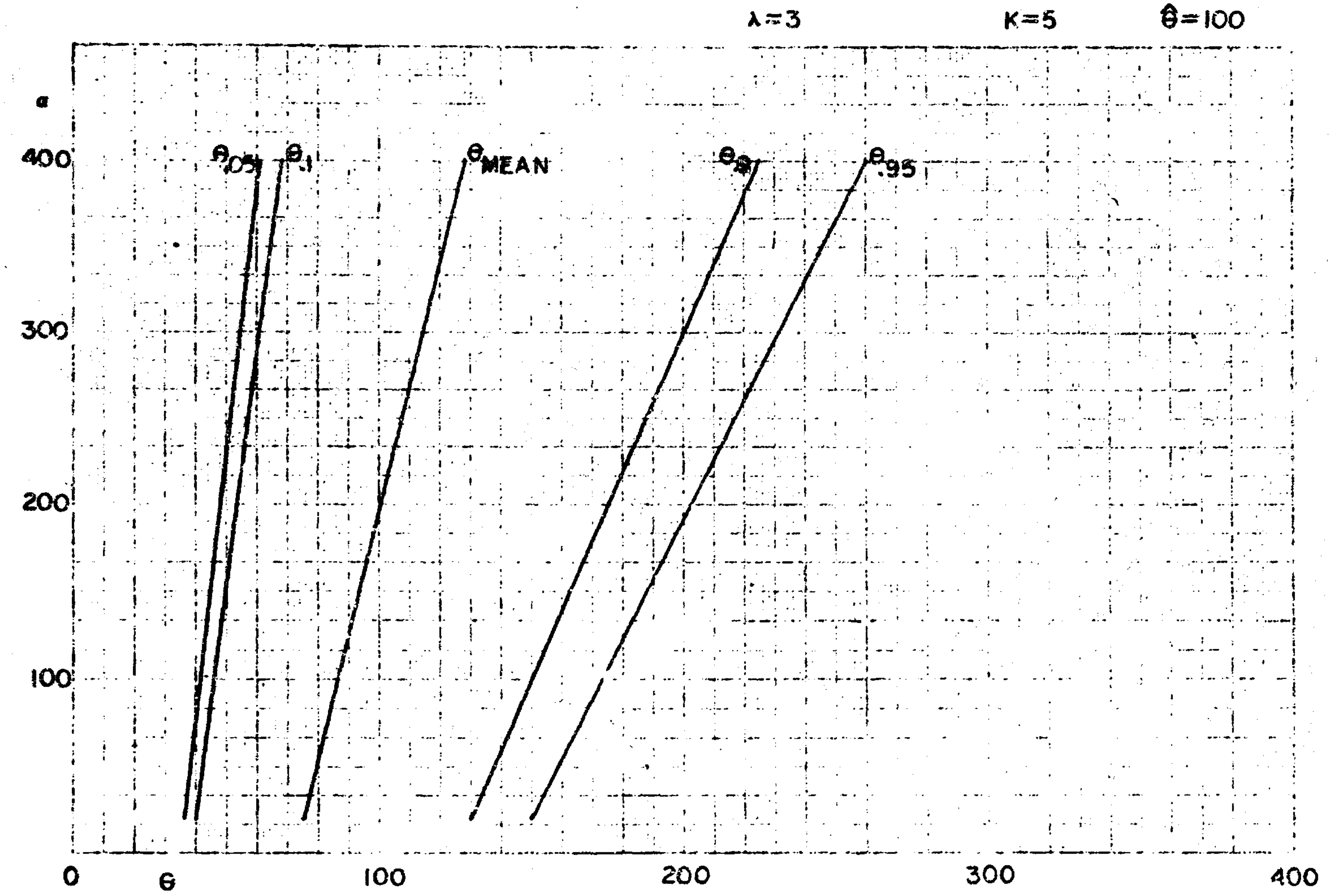


Figure 7.12 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

$\lambda=3$ $K=5$ $\hat{\theta}=50$

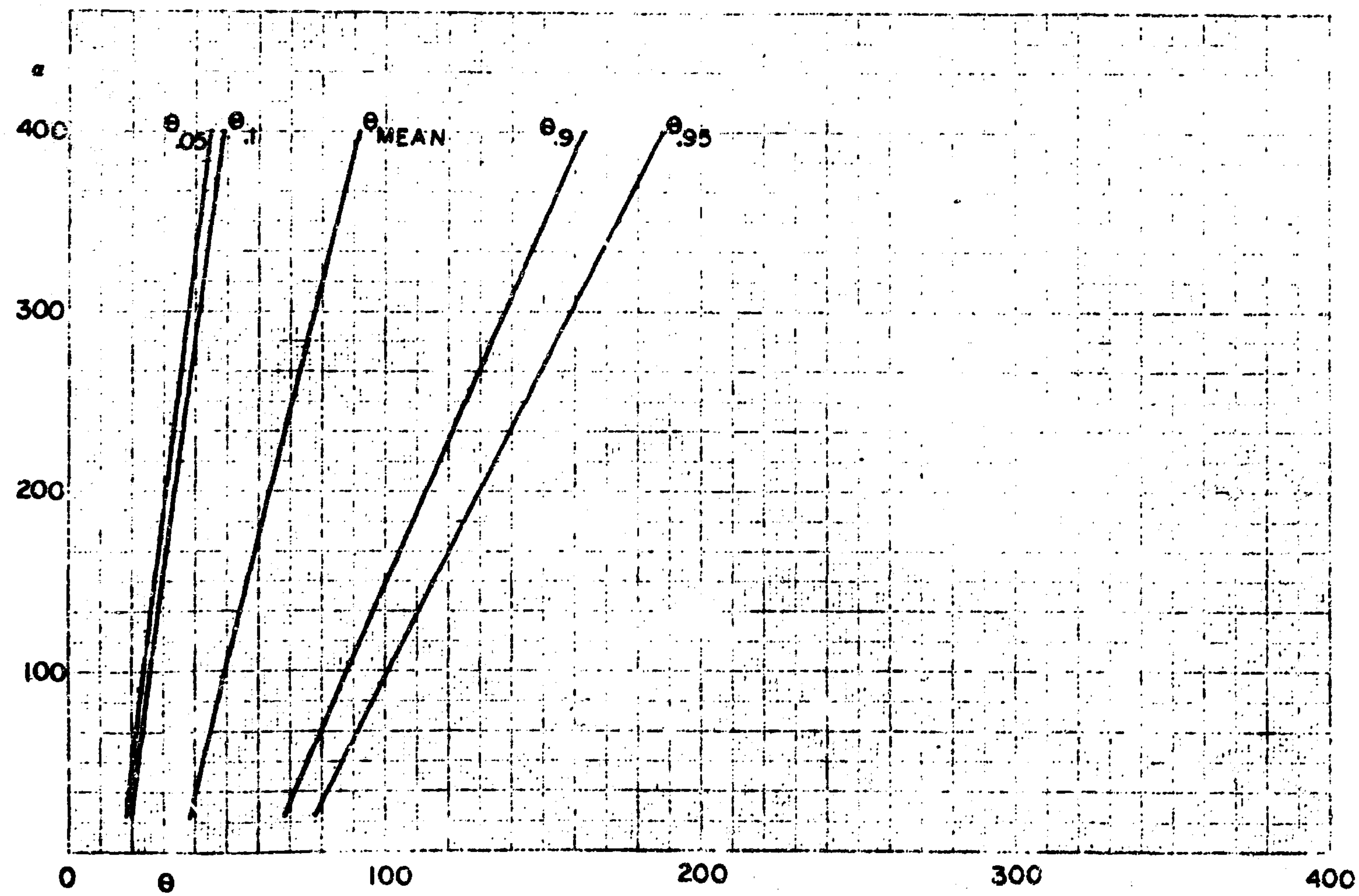


Figure 7.13 Mean and Selected Percentiles-Prior Inverted Gamma Distribution

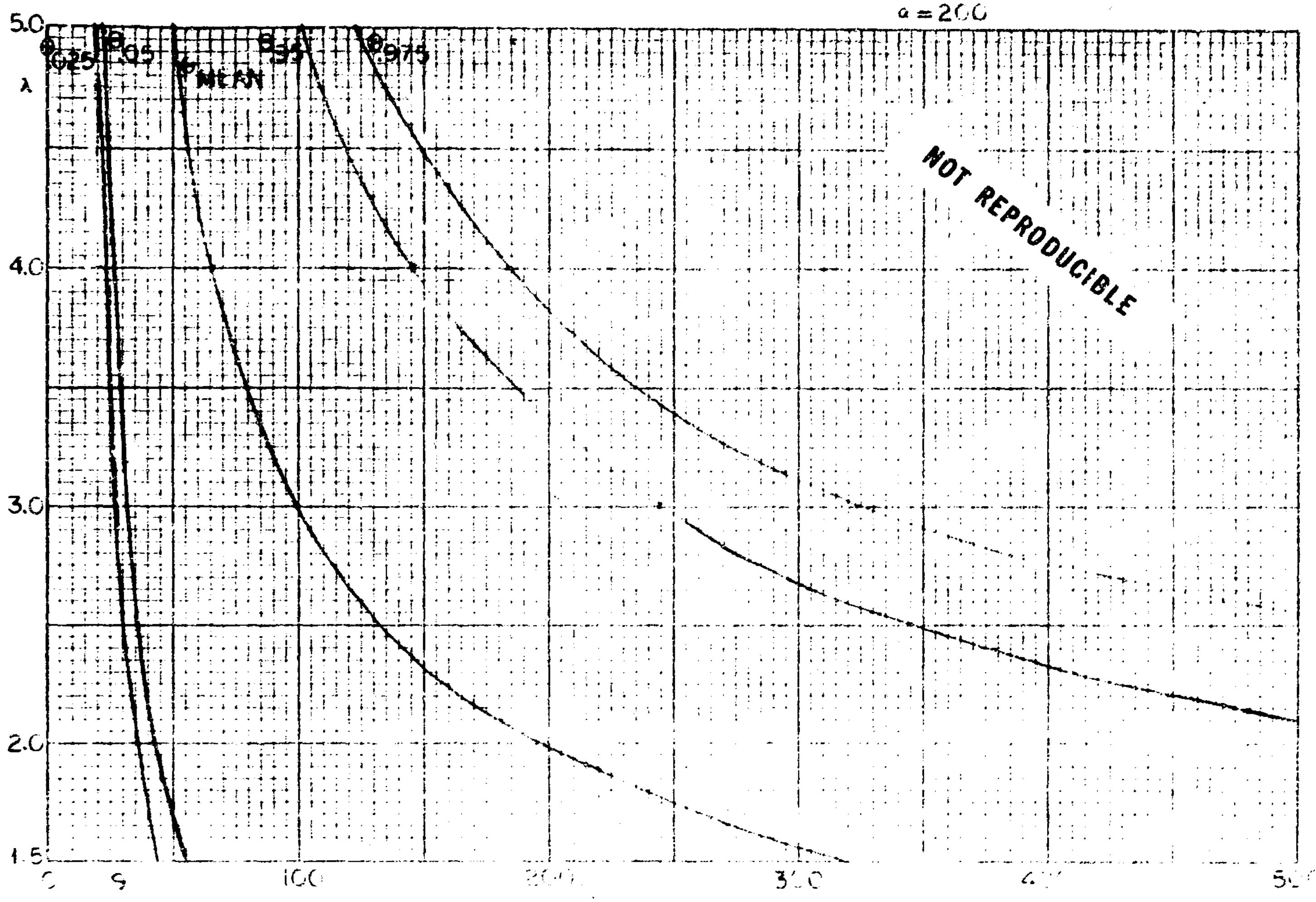


Figure 7.14 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

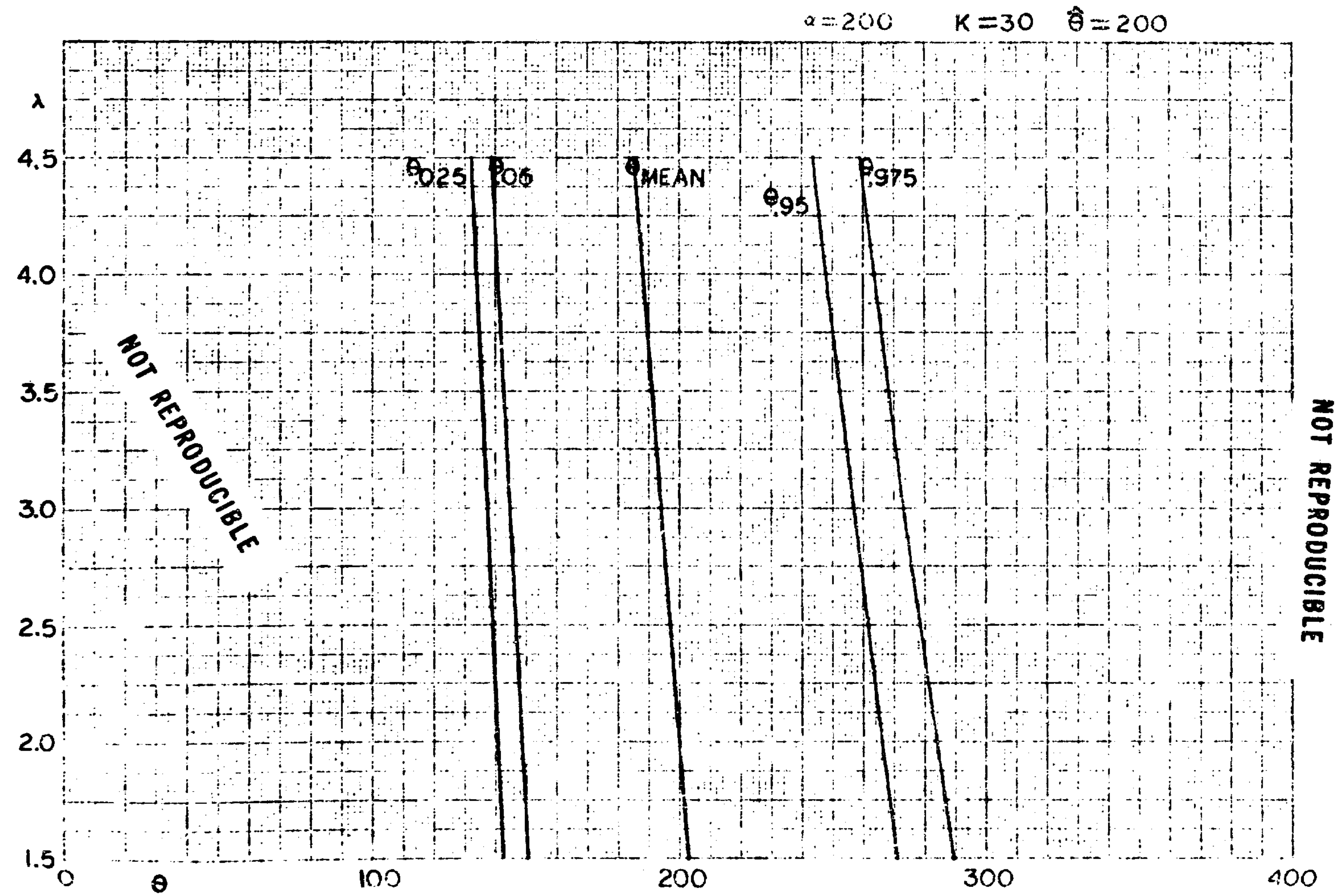


Figure 7.15 Posterior Mean and Selected Percentiles—Inverted Gamma Distribution

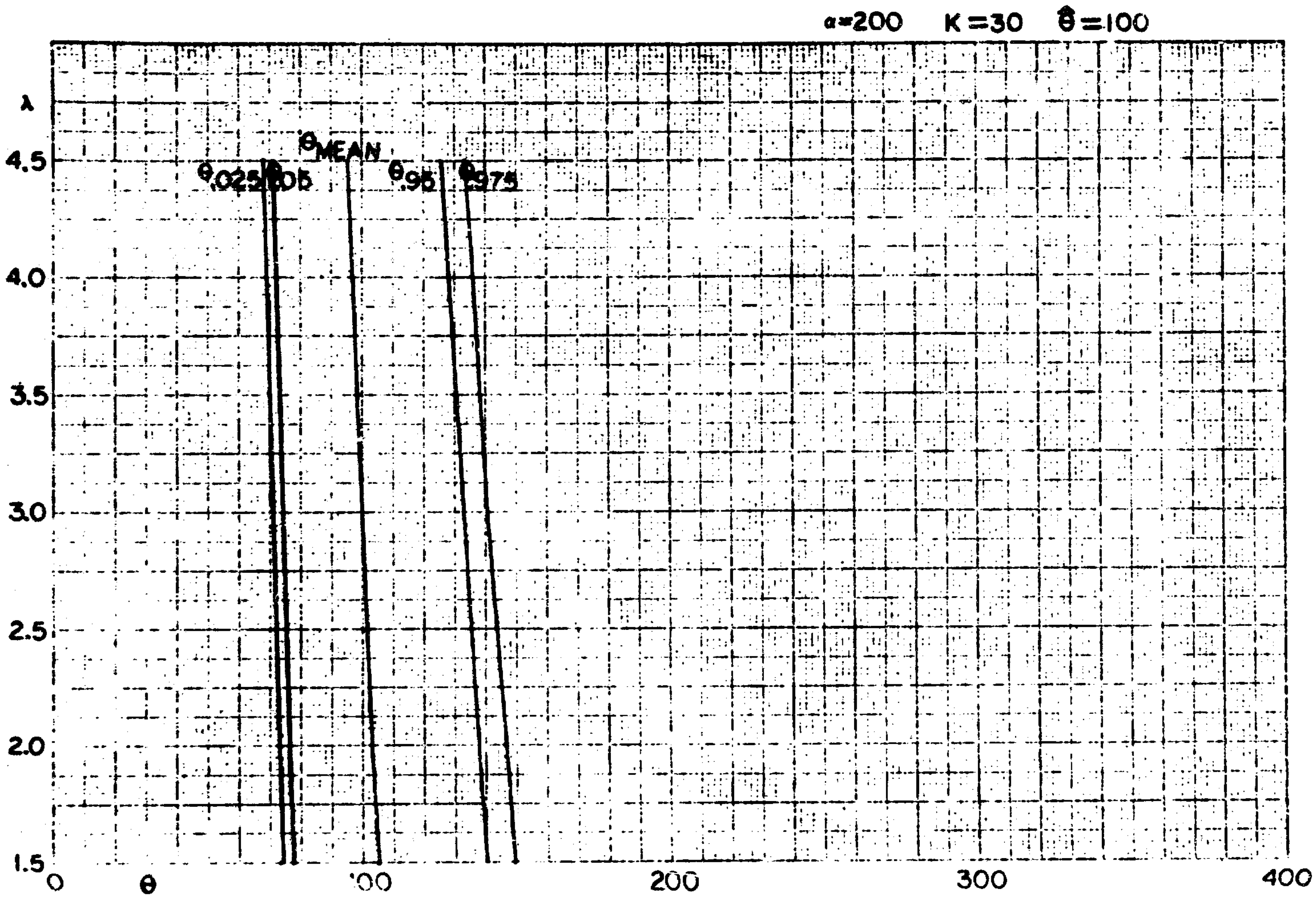


Figure 7.16 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

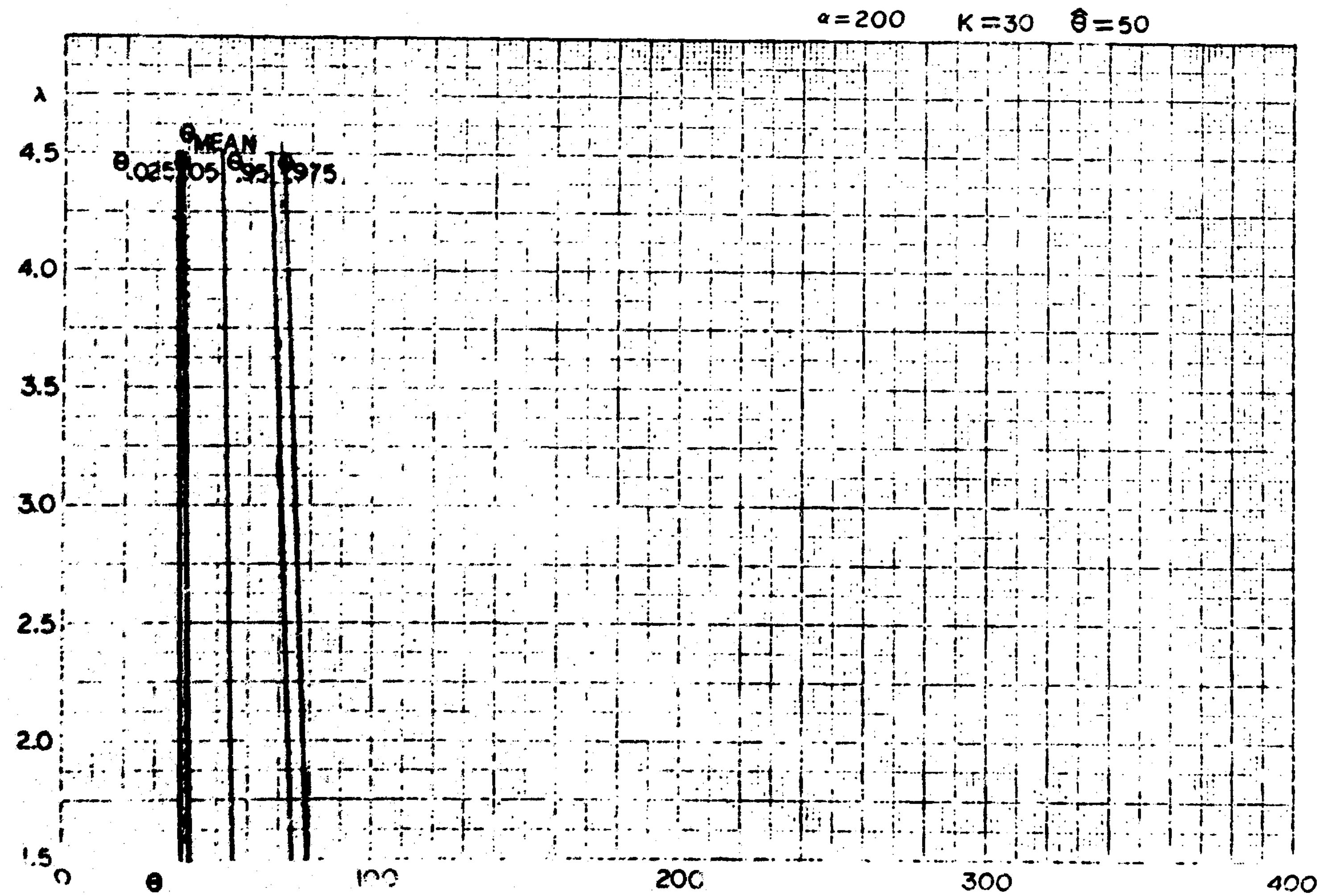


Figure 7.17 Posterior Mean and Selected Percentiles—Inverted Gamma Distribution

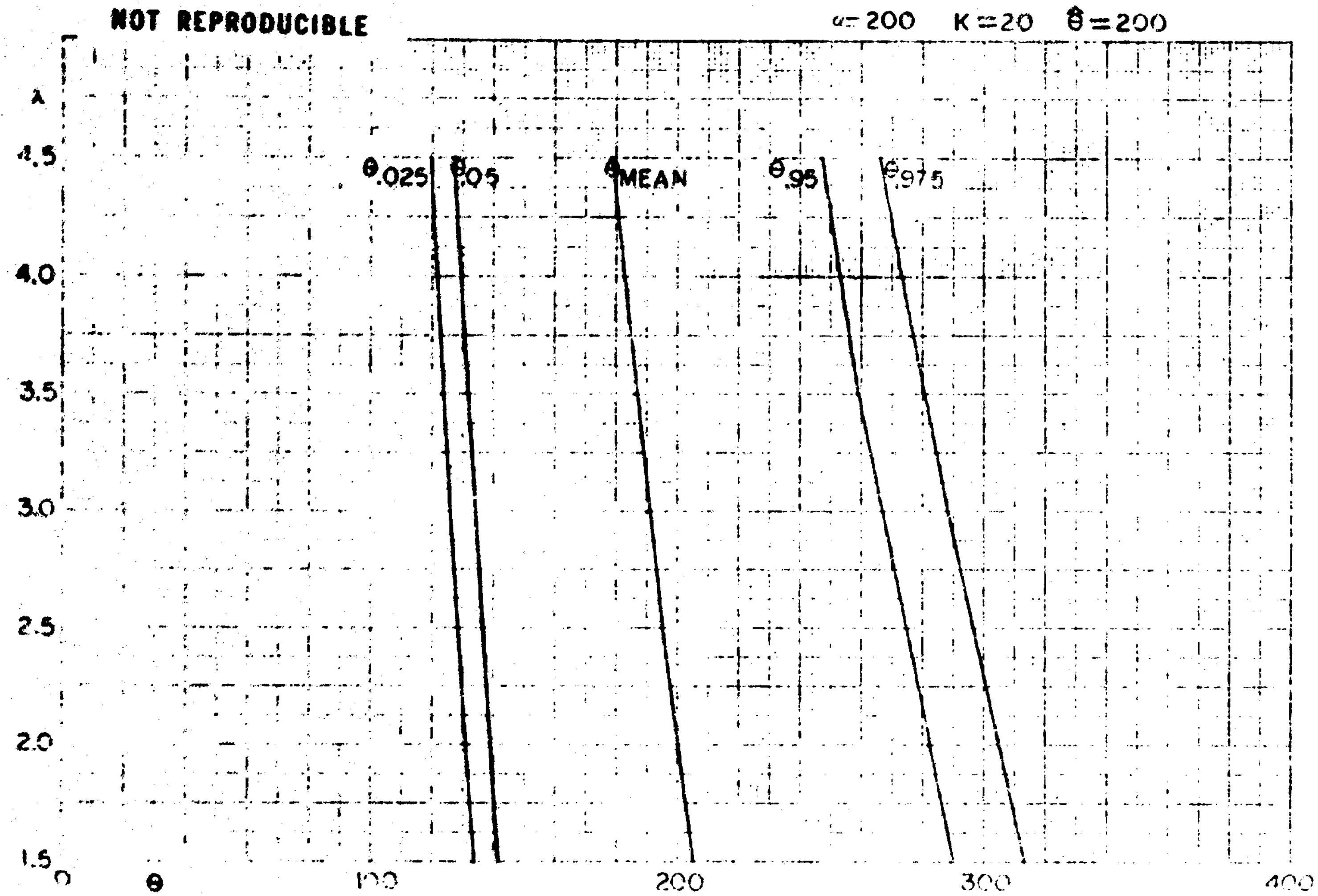


Figure 7.18 Posterior Mean and Selected Percentiles—Inverted Gamma Distribution

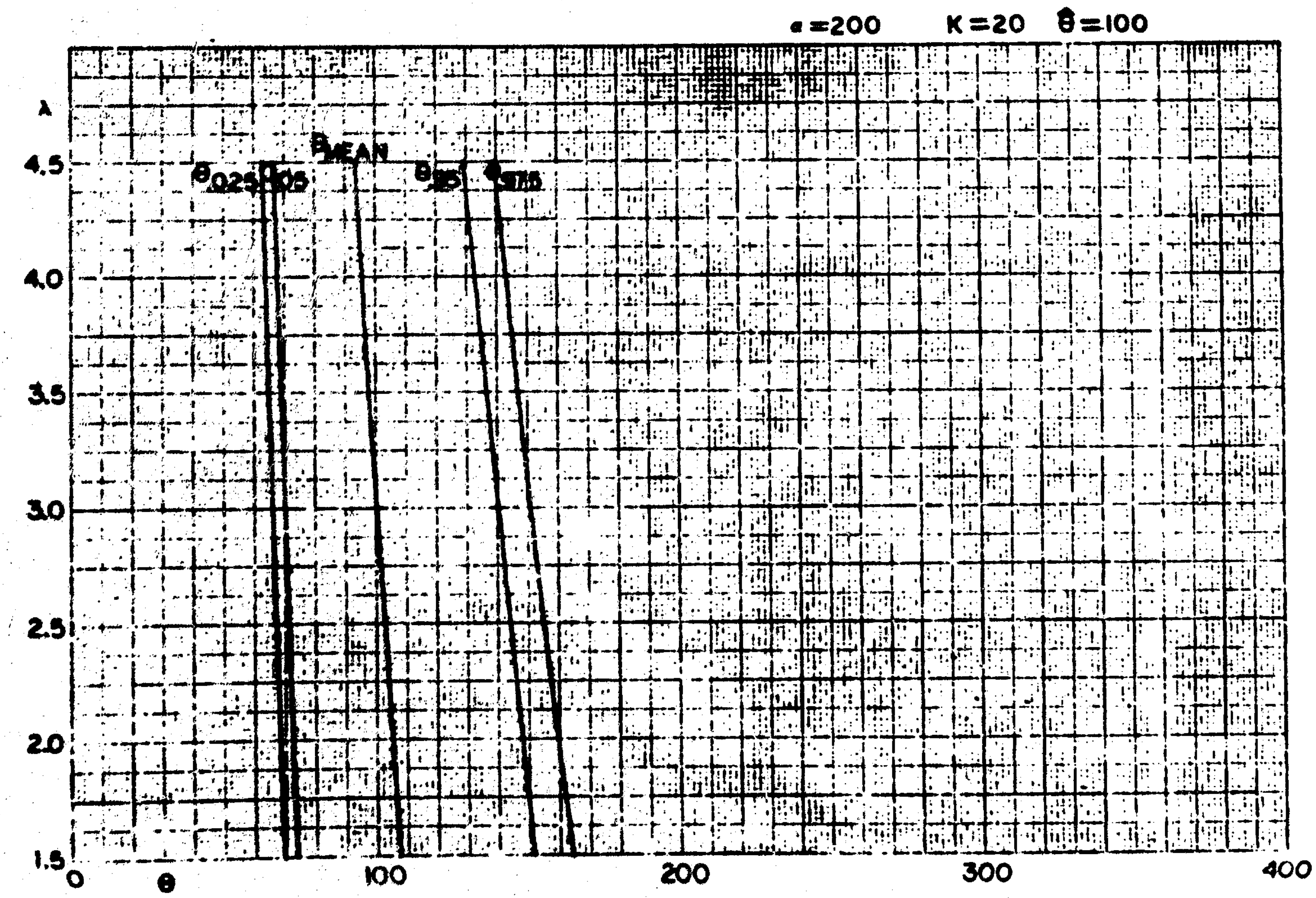


Figure 7.19 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

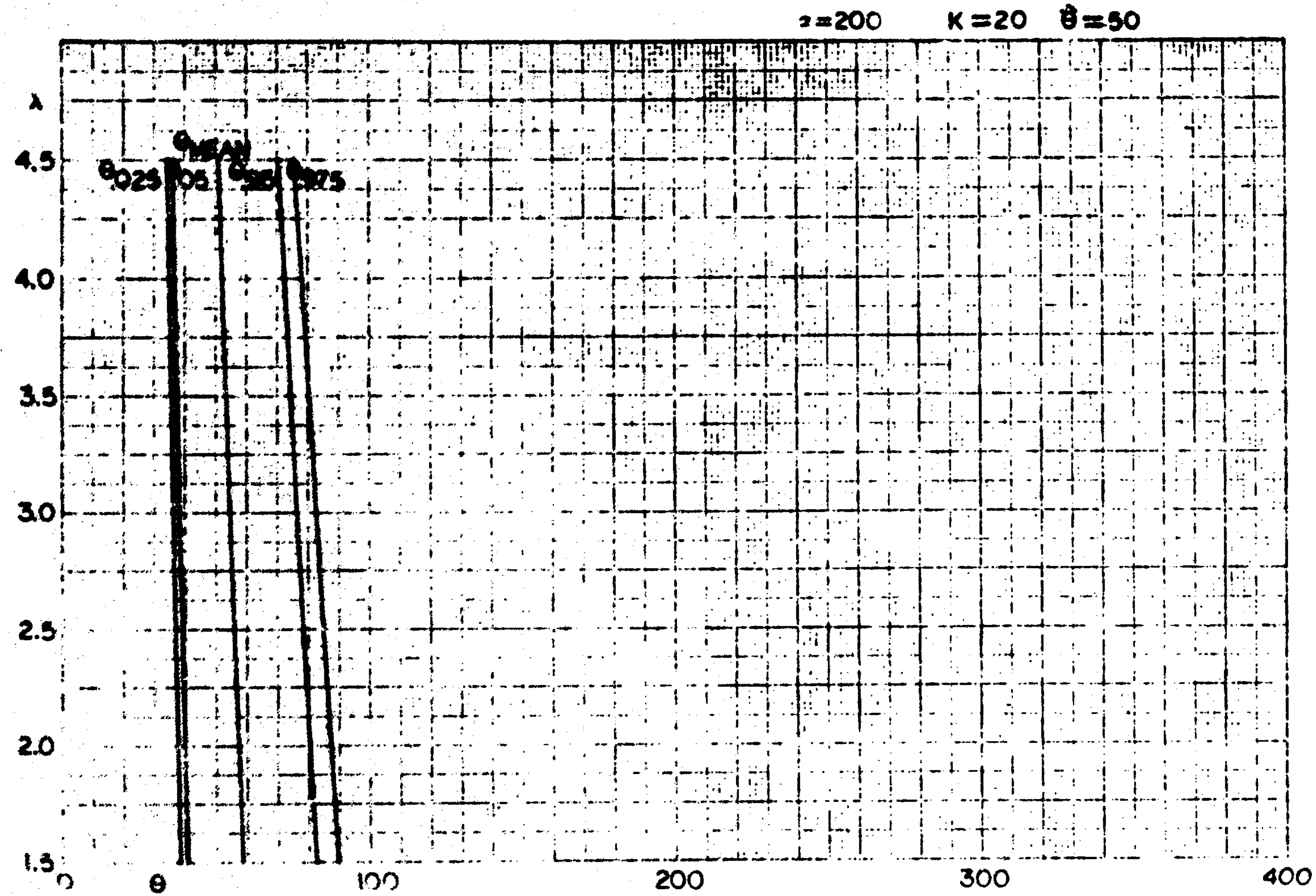


Figure 7.20 Posterior Mean and Selected Percentiles—Inverted Gamma Distribution

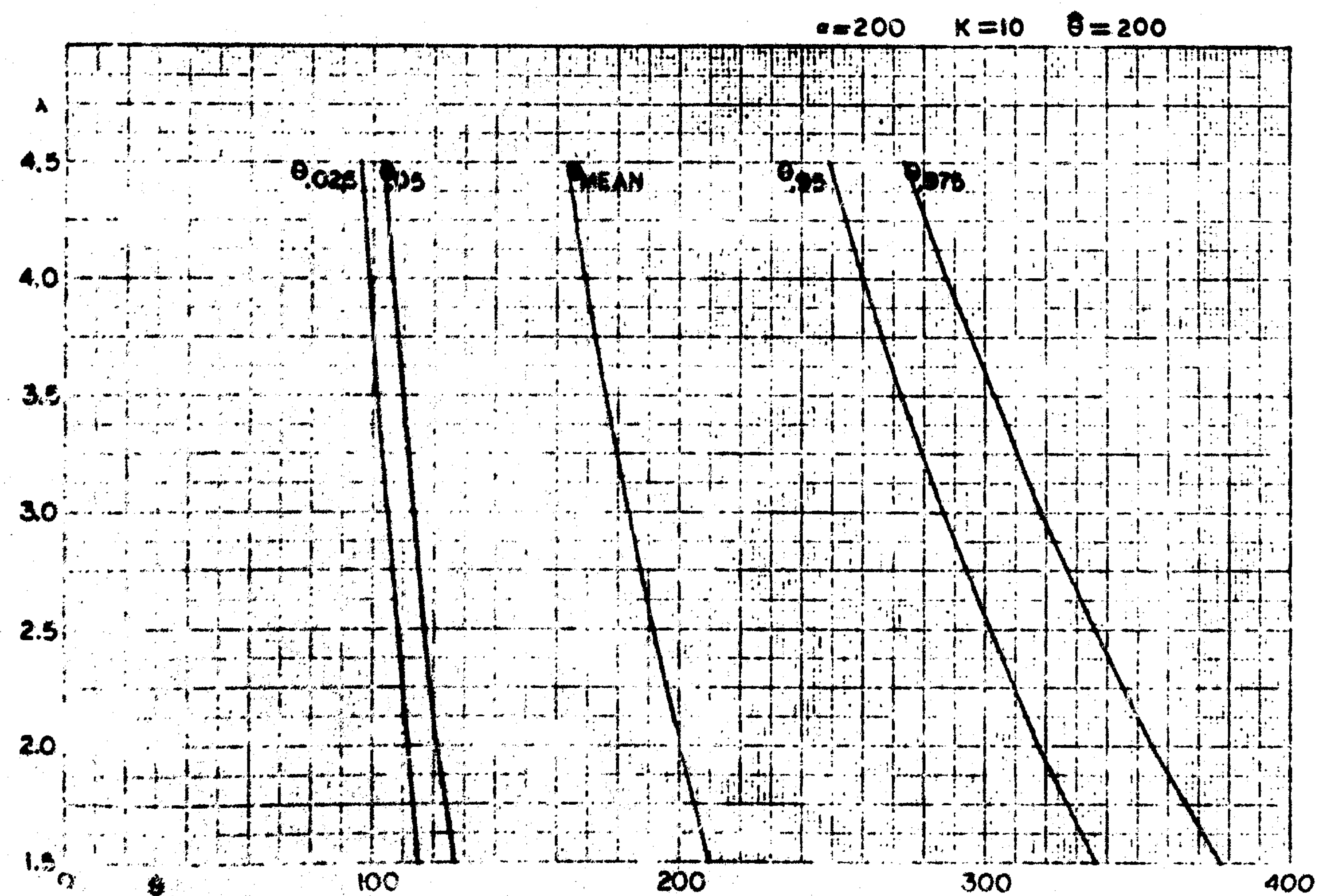


Figure 7.21 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

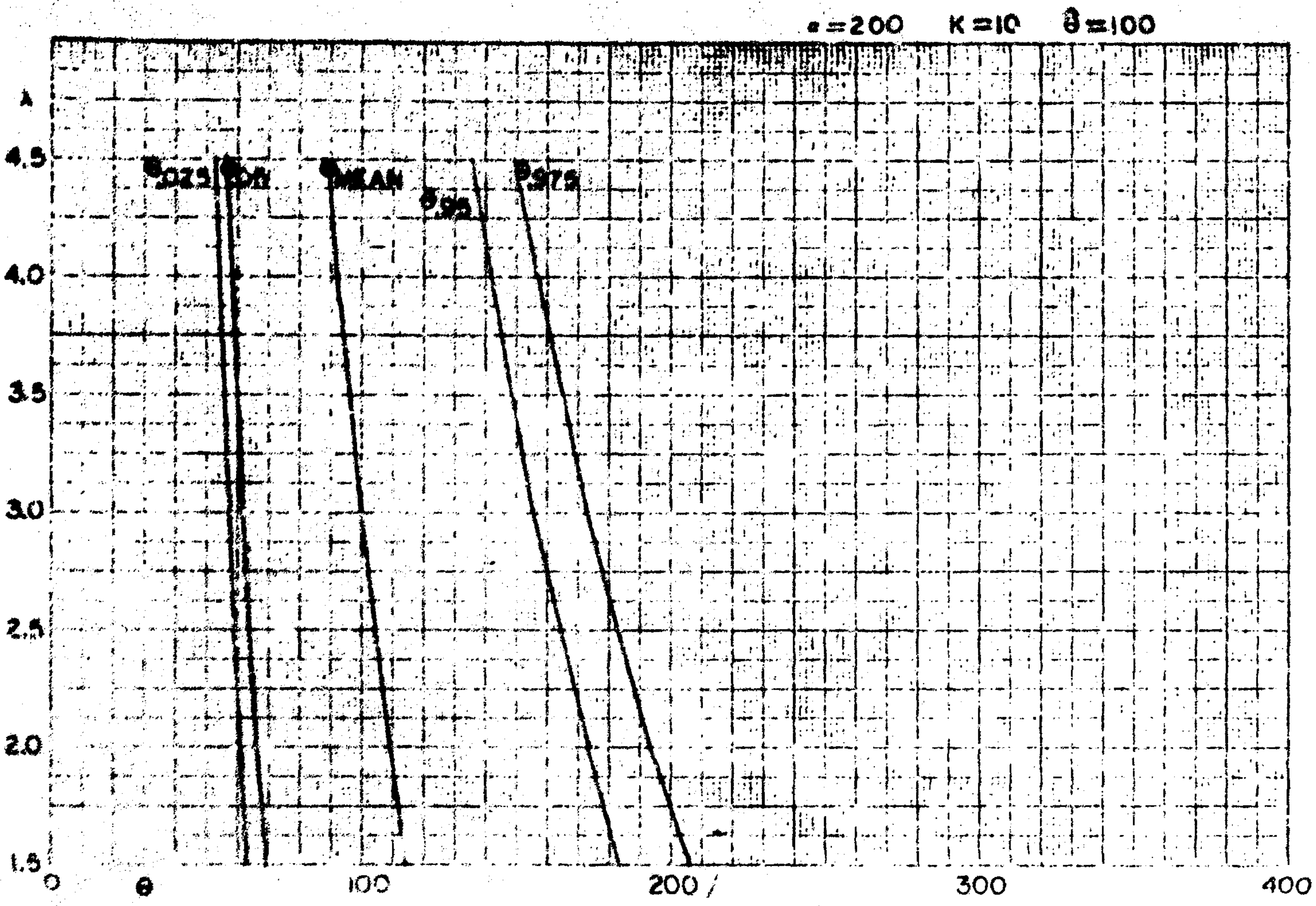


Figure 7.22 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

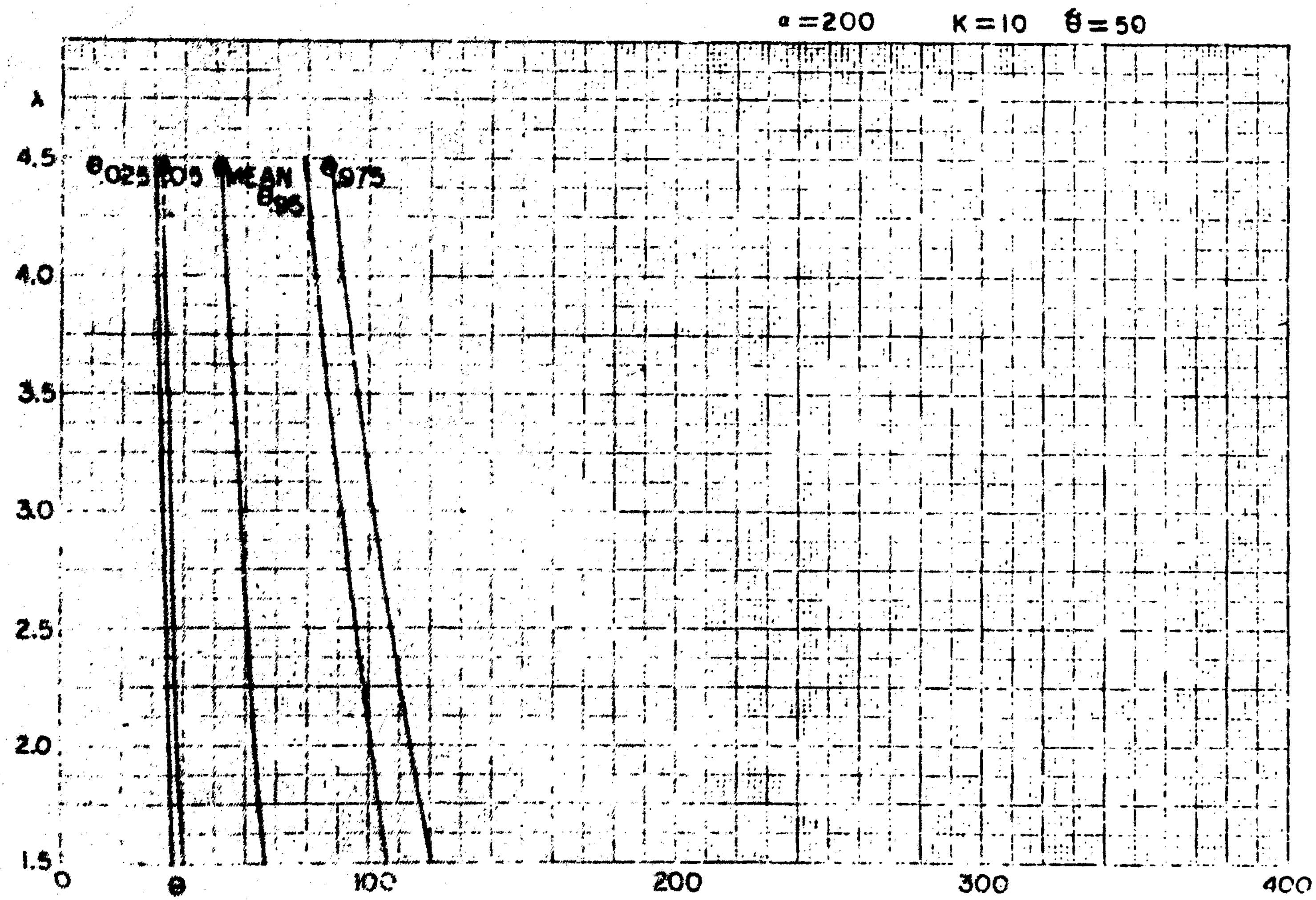


Figure 7.23 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

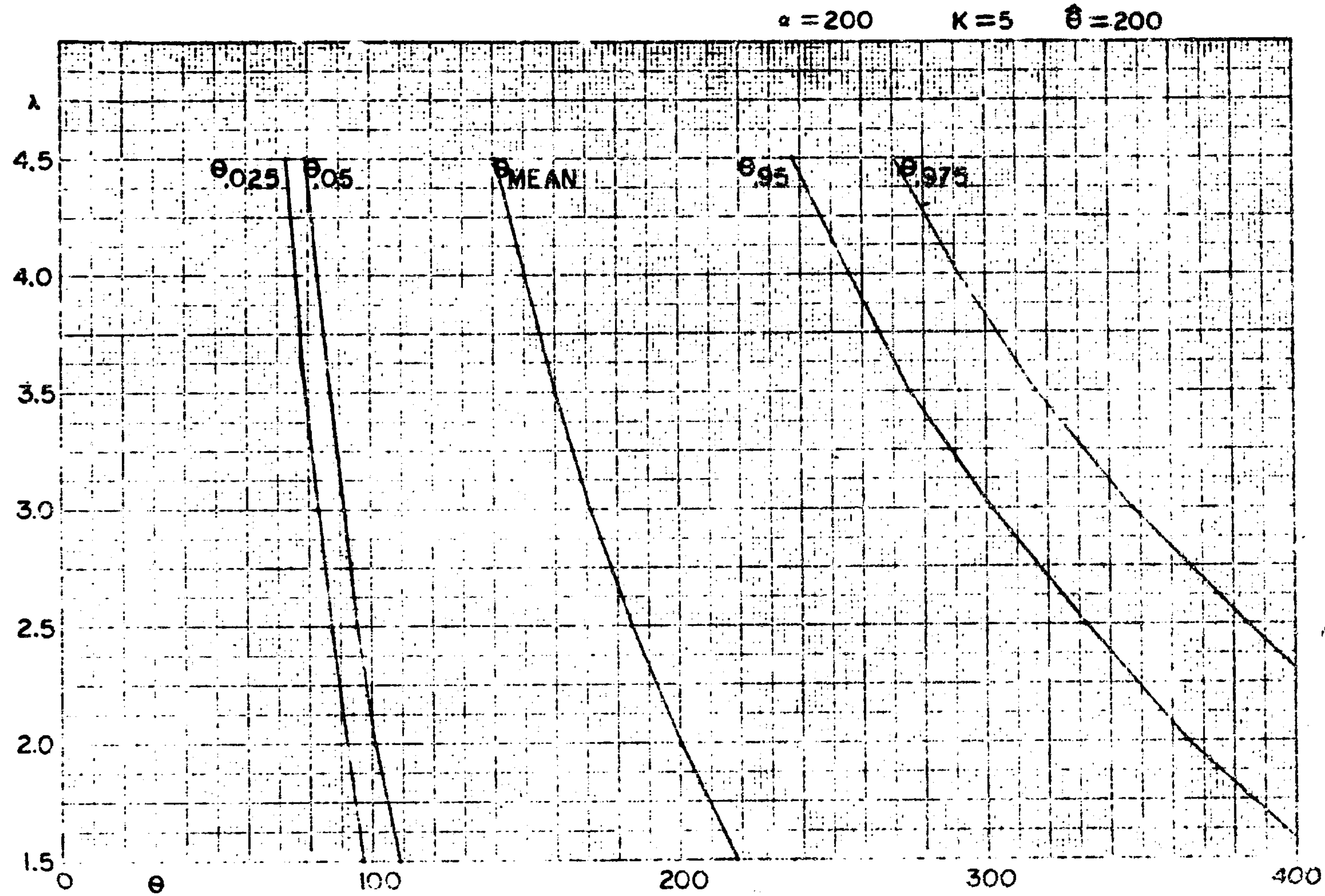


Figure 7.24 Posterior Mean and Selected Percentiles-Inverted Gamma Distribution

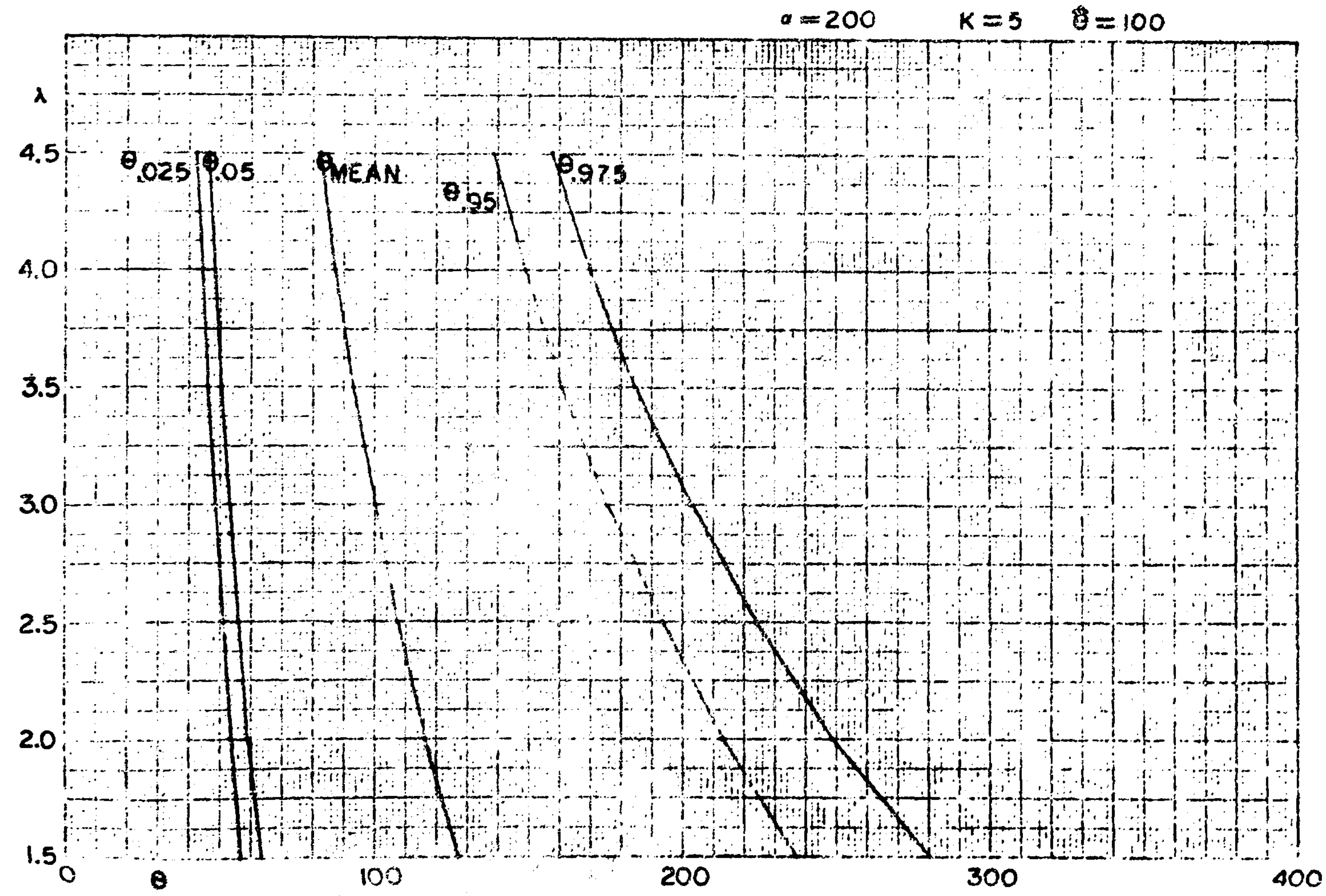
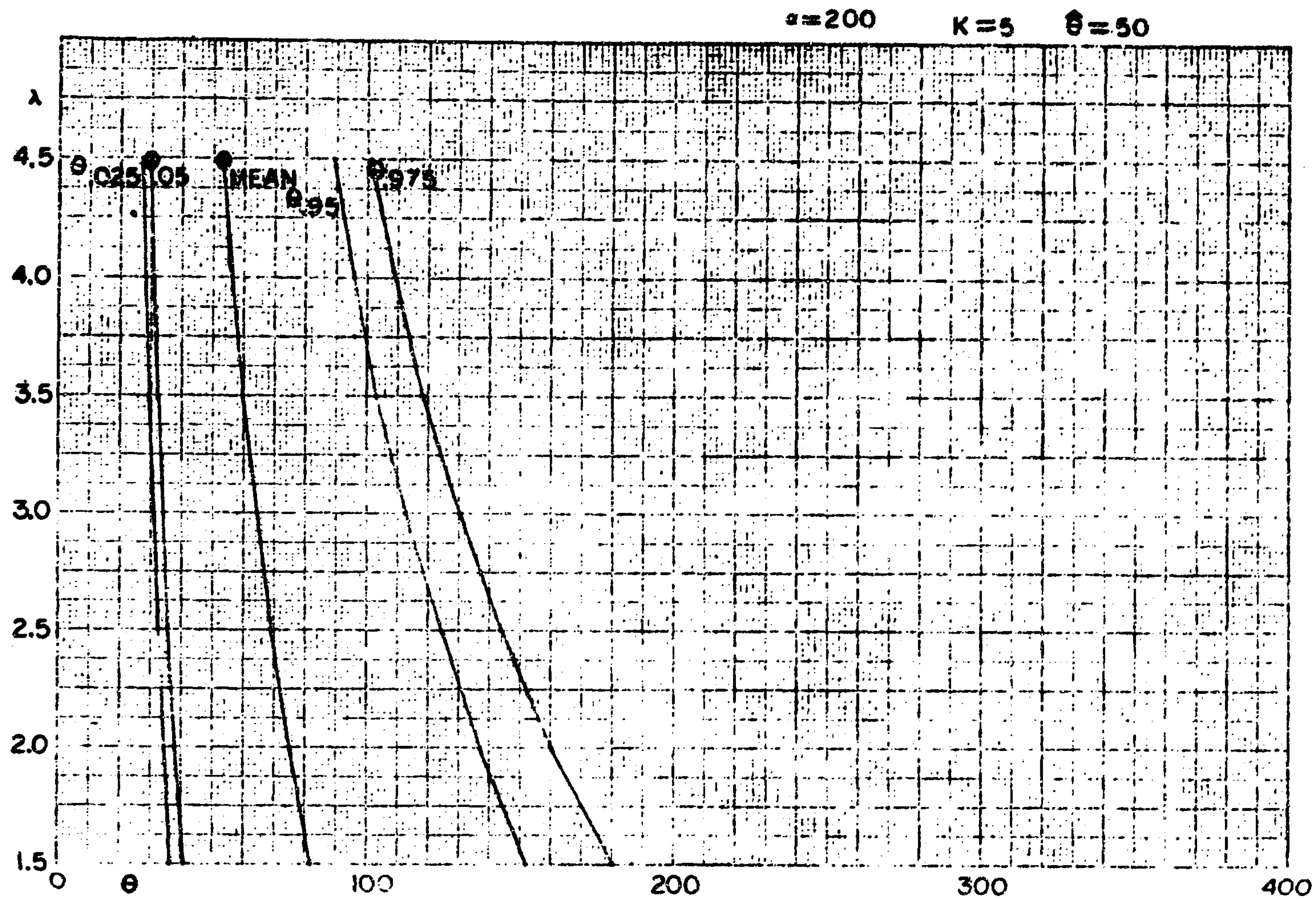


Figure 7.25 Posterior Mean and Selected Percentiles—Inverted Gamma Distribution



SECTION 8.0 CONCLUSIONS AND RECOMMENDATIONS

8.1 CONCLUSIONS

As a result of this study certain conclusions appear inescapable. We list the important ones now. First, fitting prior distributions to θ = MTBF is entirely feasible. The techniques developed in this report indicate that the mechanics of the fitting is a relatively inexpensive process. Secondly, the amount of data available, i.e., in existence today, which leads to good prior distribution fits is somewhat limited. Thus, the costs of fitting prior distributions, while not exorbitant, will be primarily incurred in acquiring, either by intensive search or by designed test, suitable data. Thirdly, the inverted gamma prior distribution, which is mathematically quite tractable, also appears to graduate data quite nicely since in the eight (8) situations where a fit could be obtained, seven of them were well described by an inverted gamma prior distribution.

In summary, the development of Bayes reliability demonstration tests appears to be quite feasible from the standpoint that the prior distribution needed can be fitted.

8.2 RECOMMENDATIONS

The following recommendations are intimately connected with the conclusions of this report. First, it is recommended that additional prior distributions be fitted. That is, prior distributions should be fitted to equipment different than the type studied in this report. This will further test the suitability of the inverted gamma prior distribution. Secondly, in view of the costs of fitting prior distribution, two areas should be studied.

i) Bayes methods of reliability demonstration which do not require a prior distribution should be studied. Such a method is empirical Bayes. It is doubtful this area will be entirely fruitful but is certainly worth a look.

ii) Methods of relating prior distributions on different, but perhaps similar, equipments so that prior distributions may be derived from one another instead of fitting new prior distributions to each one. An example would be relating two equipments by some fixed ratio of their predicted MTBF's. Finally, the receptiveness of Government and Industry to Bayes Demonstration tests should be studied for even if the methods are developed and feasible the Bayes plans will not return much on investment if they are not used.

In summary then, it is recommended that the next logical steps be taken in the development of Bayes reliability demonstration tests.

SECTION 9.0 APPENDIX

9.1 THE INVERTED GAMMA PRIOR DISTRIBUTION

At the time of this writing (late 1969) it seems impossible to exaggerate the importance of the inverted gamma distribution in reliability estimation/demonstration; particularly, when the measure of reliability is θ = MTBF. The reasons for this importance follow.

First, the inverted gamma distribution is a two parameter distribution (it can be made three parameter) and is very flexible. That is, the inverted gamma distribution can be used to graduate a wealth of empirical data. As evidence of this, we offer the fact that seven of the eight data sets analyzed resulted in good inverted gamma fits.

Secondly, and, perhaps more importantly, the marginal distribution of X (X being either time to failure or number of failures per fixed time T) is also available in closed form. This makes problems of fitting the prior distribution much more tractable.

Finally, the inverted gamma distribution is the natural conjugate for the Poisson process (i.e., for the exponential and Poisson conditional distributions). This means that the posterior distribution is of the same form as the prior distribution (inverted gamma) and hence, the posterior distribution is available in closed form. This makes computations for demonstration tests and other analyses much more tractable.

The gamma distribution in the random variable X is given by

$$g(x) = \frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} \quad \alpha,\lambda,x > 0. \tag{9.1.1*}$$

The change of variable $\theta = 1/x$ leads to

$$g(\theta) = \frac{\alpha^\lambda}{\Gamma(\lambda)} \theta^{-(\lambda+1)} e^{-\alpha/\theta} \tag{9.1.2}$$

Thus, the gamma and inverted gamma are related by a reciprocal transformation. This is particularly convenient since for the Poisson process the MTBF is the reciprocal of the failure rate. Hence, a gamma prior distribution on the failure rate implies an inverted gamma prior distribution on the MTBF and conversely.

---

*$\Gamma(\lambda)$ is notation for the gamma function: $\Gamma(\lambda) = \int_0^\infty y^{\lambda-1} e^{-y}\, dy$.

The characteristic function of the inverted gamma is useless for finding the moments but they may be found directly. The K^{th} moment is given by

$$E(\theta^K) = \int_0^\infty \theta^K g(\theta)\,d\theta$$

$$= \frac{\alpha^\lambda}{\Gamma(\lambda)} \int_0^\infty \theta^{-(\lambda+1-K)} e^{-\alpha/\theta}\,d\theta \qquad (9.1.3)$$

The change of variable $y = \alpha/\theta$ leads to

$$E(\theta^K) = \frac{\alpha^K}{\Gamma(\lambda)} \int_0^\infty y^{(\lambda-K)-1} e^{-y} dy \qquad (9.1.4)$$

For $(\lambda-K) > 0$ the integral in (9.1.4) converges and because of the recurrence $\Gamma(v)v = \Gamma(v+1)$

$$E(\theta^K) = \frac{\alpha^K}{\prod_{i=1}^{K} (\lambda-i)} \qquad (9.1.5)$$

For $(\lambda-K) \leq 0$, i.e., $\lambda \leq K$, the integral in (9.1.4) is infinite and $E(\theta^K)$ does not exist.

If $[\lambda]$ = the largest integer smaller than λ (e.g., $[4.3] = 4$, $[4.0] = 3$) then all moments up to and including $[\lambda]$ exist and no moments beyond $[\lambda]$ exist. Using (9.1.5):

$$E(\theta) = \frac{\alpha}{\lambda-1} \qquad \lambda > 1$$

$$E(\theta^2) = \frac{\alpha^2}{(\lambda-1)(\lambda-2)} \qquad \lambda > 2 \qquad (9.1.6)$$

and

$$\sigma_\theta^2 = \frac{\alpha^2}{(\lambda-1)^2(\lambda-2)}$$

EXAMPLE: Suppose, given an inverted gamma distribution with $\alpha = 100$ $\lambda = 2.5$. Then $[\lambda] = 2$ and only the first two moments exist. They are

$$E(\theta) = \frac{100}{1.5} = 66.7$$

$$E(\theta^2) = \frac{10{,}000}{.75} = 13{,}333.3.$$

It is of no consequence to this study that the moments do not all exist. Other than this the inverted gamma is well-behaved and even if the mean does not exist, the mode, median and all quantiles do. The mode is given by

$$\text{Mode} = \frac{\alpha}{\lambda+1}\,. \tag{9.1.7}$$

9.2 THE MARGINAL DISTRIBUTIONS

In this section the marginal distributions (for an inverted gamma prior distribution) will be derived for two conditional distributions.

i) The Poisson distribution when number of failures occurring in fixed time T is the observed random variable.

ii) The gamma distribution when the time to failure is exponential and hence, the observed random variable is sample MTBF $\hat{\theta}$.

For case i) above

$$f(x) = \int_0^\infty f(x|\theta)\, g(\theta)\, d\theta$$

$$= \int_0^\infty \left[\frac{e^{-T/\theta}(T/\theta)^x}{x!}\right]\left[\frac{\alpha^\lambda}{\Gamma(\lambda)}\,\theta^{-(\lambda+1)}\,e^{-\alpha/\theta}\right] d\theta$$

$$= \frac{\alpha^\lambda}{\Gamma(\lambda)}\,\frac{T^x}{x!}\int_0^\infty \frac{e^{-(T+\alpha)/\theta}}{\theta^{\lambda+x+1}}\,d\theta.$$

With the change of variable $Z = (T+\alpha)/\theta$

$$f(x) = \frac{T^x}{x!}\,\frac{\alpha^\lambda}{\Gamma(\lambda)}\,\frac{1}{(T+\alpha)^{\lambda+x}}\int_0^\infty e^{-Z}\,Z^{\lambda+x-1}\,dZ.$$

The integral term is $\Gamma(\lambda+x)$ so that

$$f(x) = \frac{\Gamma(\lambda+x)}{\Gamma(\lambda)x!}\left(\frac{T}{T+\alpha}\right)^x\left(\frac{\alpha}{T+\alpha}\right)^\lambda \tag{9.2.1}$$

Setting $p = \frac{\alpha}{T+\alpha}$ and $q = \frac{T}{T+\alpha}$ it is noted that $p+q = 1$ and

$$f(x) = \frac{\Gamma(\lambda+x)}{\Gamma(\lambda)x!} p^{\lambda} q^{x} \qquad x = 0,1, \ldots ,$$

is the well-known negative binomial distribution with

$$\text{Mean} = E(x) = \frac{\lambda T}{\alpha}$$

$$\text{Variance} = \sigma_x^2 = \frac{\lambda T(T+\alpha)}{\alpha^2}$$

It should be noted that if we had started with a gamma distribution on failure rate and a Poisson conditional with mean: failure rate times T instead of T/θ the same negative binomial distribution as (9.2.1) would have resulted.

In case ii) above, it is easy to show that if time to failure is exponential, then the conditional distribution of the sample MTBF, $\hat{\theta}$, given the MTBF, θ, is gamma, i.e.,

$$f_K(\hat{\theta}|\theta) = \left(\frac{K}{\theta}\right)^K \frac{1}{\Gamma(K)} (\hat{\theta})^{K-1} e^{\frac{-K\hat{\theta}}{\theta}} .$$

Here, K is the observed number of failures and θ is the true but unknown MTBF. Hence, the marginal distribution of $\hat{\theta}$ is

$$f_K(\hat{\theta}) = \int_0^\infty f_K(\hat{\theta}|\theta)\, g(\theta)\, d\theta$$

$$= \frac{\alpha^{\lambda} K^{K}}{\Gamma(\lambda)\Gamma(K)} \int_0^\infty e^{-1/\theta(\alpha+K\hat{\theta})} \frac{\hat{\theta}^{K-1}}{\theta^{K+\lambda+1}}\, d\theta .$$

The change of variable $Z = \frac{\alpha+K\hat{\theta}}{\theta}$ leads to

$$f_K(\hat{\theta}) = \frac{\Gamma(K+\lambda)}{\Gamma(K)\Gamma(\lambda)} \left(\frac{\alpha}{\alpha+K\hat{\theta}}\right)^{\lambda} \left(\frac{K\hat{\theta}}{\alpha+K\hat{\theta}}\right)^{K-1} \left(\frac{K}{\alpha+K\hat{\theta}}\right) \hat{\theta} > 0 . \qquad (9.2.2)$$

which is the not so well-known inverted Beta distribution.

In any event

$$E(\hat{\theta}) = \frac{\alpha}{\lambda-1} \qquad \lambda > 1$$

$$E(\hat{\theta}^2) = \frac{\alpha^2(K+1)}{K(\lambda-1)(\lambda-2)} \qquad \lambda > 2.$$

9.3 TRANSFORMATIONS ON THE INVERTED GAMMA DISTRIBUTION

In Section 6.0 (Data Combination) a certain method of relating prior distributions on "different" equipments was discussed. The idea depends on a property of the inverted gamma distribution (many other distributions have this same property) which will be shown here:

If

$$g(\theta) = \frac{\alpha^{\lambda}}{\Gamma(\lambda)} \theta^{-(\lambda+1)} e^{-\alpha/\theta}, \qquad (9.3.1)$$

i.e., if θ is inverted gamma with scale parameter α and shape parameter λ then the random variable $y = \beta\theta$, $0 < \beta < \infty$, is again inverted gamma with scale parameter $(\alpha\beta)$ and shape parameter λ. That is,

$$g(y) = \frac{(\alpha\beta)^{\lambda}}{\Gamma(\lambda)} y^{-(\lambda+1)} e^{-\alpha\beta/y} \qquad (9.3.2)$$

This result follows by substituting $\theta = \frac{y}{\beta}$ in (9.3.1) and multiplying by the differential element $d\theta = dy/\beta$.

A change of the form $\theta = \frac{y}{\beta} - c$, $0 < c < \infty$, has the effect of introducing a guarantee time βc and changing the scale parameter to $\alpha\beta$.

A change of the form $\theta = y^{\lambda}$ means that y is not of the inverted gamma family.

9.4 IDENTIFIABILITY OF THE PRIOR DISTRIBUTION

The central problem in estimating the prior distribution of MTBF, say $g(\theta)$, is that random samples (to construct an estimate of $g(\theta)$) are not available from $g(\theta)$. Thus, the marginal distribution

$$f(x) = \int_0^{\infty} f(x|\theta) g(\theta) d\theta \qquad (9.4.1)$$

must be used to estimate $g(\theta)$. Suppose now that the equation

$$f(x) = \int_0^\infty f(x|\theta)\, g_1(\theta)\, d\theta = \int_0^\infty f(x|\theta)\, g_2(\theta)\, d\theta \qquad (9.4.2)$$

implies that $g_1(\theta) = g_2(\theta)$. In this situation $f(x)$ is called identifiable with respect to $f(x|\theta)$. Identifiability is a crucial property for this study, since if (9.4.2) could hold and $g_1(\theta) \neq g_2(\theta)$, then two different prior distributions could lead to the same marginal distribution and since the marginal distribution is used to make inferences about $g(\theta)$ it would be impossible to tell which (in the case of non identifiability) $g(\theta)$ obtained.

We will show the identifiability of the marginal distribution $f(x)$ with respect to three conditional densities (the Poisson, exponential, and gamma).

i) $f(x|\theta) = \dfrac{e^{-T/\theta}(T/\theta)^x}{x!}$. This is the case where the number of failures occurring in fixed time T is the observed random variable and the operating process is the Poisson process.

ii) $f(x|\theta) = 1/\theta\, e^{-x/\theta}$. Here the random variable X is time to failure.

iii) $f(\hat{\theta}|\theta) = \dfrac{(K/\theta)^K}{\Gamma(K)} (\hat{\theta})^{K-1} e^{-\frac{K\hat{\theta}}{\theta}}$. This is the gamma distribution of sample MTBF when times to failure are exponential.

A result of Teicher (Reference 7) is used to show that $f(x)$ is identifiable w.r.t. i) above, i.e., that different prior distributions cannot lead to the same marginal distribution. The result is that a family of densities which is additively closed (a.c.) is identifiable. Additively closed means

$$f(x|\alpha) * f(x|\beta) = f(x|\alpha + \beta) \qquad (9.4.3)$$

The * denotes convolution. Since it is easy to show that the sum of two Poisson variates is again a Poisson variate with parameter the sum of the two parameters $f(x)$ is identifiable in case i).

For case ii), we note that

$$f(x) = \int_0^\infty 1/\theta\, e^{-x/\theta} g(\theta)\, d\theta \qquad (9.4.4)$$

Making the change of variable $\lambda = 1/\theta$

$$f(x) = \int_0^\infty \lambda e^{-\lambda x} g(\lambda) d\lambda \tag{9.4.5}$$

Thus, $f(x)$ divided by λ is the Laplace transform of $g(\lambda)$. If $g(\lambda)$ is a class of continuous densities (e.g., the gamma family) then $g(\theta)$is a continuous class (e.g., inverted gamma) and by the uniqueness theorem for Laplace transforms $f(x)$ is identifiable.

Finally, for case iii) we again use a result of Teicher (Reference 7). That is, for fixed K iii) is a scale parameter family generated by $f(x|n,1)$ and a uniqueness theorem for Fourier transforms gives the desired result.

9.5 LIMITING BEHAVIOR OF THE MARGINAL DENSITY $f_K(\hat{\theta})$

When $\hat{\theta}$ is the sample MTBF based on K failures, then under the conditions **given below** the marginal distribution $f_K(\hat{\theta})$ converges to the prior distribution $g(\theta)$ as $K \to \infty$. This is shown in the result and proof 9.5.1 below. Figure 9.5.1 shows a particular example: the prior density $g(\theta)$ is plotted along with the corresponding inverted Beta marginal densities $f_K(\hat{\theta})$ for $K = 1,2,3,4,5,10,20,50$. Clearly, when K is small, it is erroneous to fit sample values of $\hat{\theta}$ directly to $g(\theta)$. When K is large, however, (any $K \geq 20$), the error is not too bad, and sample $\hat{\theta}$'s based on large K's can be fitted directly to $g(\theta)$. Thus, when K is "large" enough, a fairly accurate method exists for fitting the prior which avoids the problems involved in using a mixed population model for the marginal distribution.

9.5.1 A RESULT AND PROOF

Consider a density $g(\theta)$ defined on $(0,\infty)$ such that sup $g(\theta)$ is finite, and consider a sequence of R.V.'s $\{T_K\}$ with densities given by

$$h_K(t) = \int_0^\infty g(\theta) f_K(t|\theta) d\theta, \; K = 1,2, \; \ldots \text{ where the}$$

sequence $\{f_K\}$ has the property that $\forall$ fixed $t_o > 0$,

$$\lim_{K \to \infty} \int_0^{t_o} f_K(t|\theta) dt = \begin{cases} 1 & \text{if } \theta < t_o \\ 0 & \text{if } \theta > t_o \end{cases}$$

[The above condition is satisfied in the usual Bayes set-up, where T_K is (say) the sample mean $\hat{\theta}_K$, based on a sample of size K from (say) an exponential distribution with mean θ. For then $E(\hat{\theta}_K|\theta) = \theta$, all K, and $\lim_{K\to\infty} Var(\hat{\theta}_K|\theta) = 0$, from which the above condition can easily be shown to hold.]

Then we have the following

<u>RESULT</u>: The d.f. $H_K(T_K) = P(T_K < t)$ converges pointwise to the d.f. $G(\theta) = P(\theta < t)$, i.e., $\forall$ fixed $t_o > 0$, such that $\lim_{K\to\infty} H_K(t_o) = G(t_o)$.

<u>PROOF</u>: Fix t_o. By definition,

$$H_K(t_o) = \int_0^{t_o}\int_0^{\infty} g(\theta)f_K(t|\theta)d\theta\, dt$$

$$= \int_0^{\infty} g(\theta)\int_0^{t_o} f_K(t|\theta)dt\, d\theta, \text{ assuming the appropriate}$$

measurability conditions for Fubini's Theorem hold.

Now set $Q_K(\theta) = g(\theta)\int_0^{t_o} f_K(t|\theta)\, dt$, so that $H_K(t_o) = \int_0^{\infty} Q_K(\theta)d\theta$.

Since $\forall$ θ and K, $Q_K(\theta)$ is bounded above by sup g, we apply the dominated convergence theorem and get

$$\lim_{K\to\infty} H_K(t_o) = \lim_{K\to\infty}\int_0^{\infty} Q_K(\theta)d\theta = \int_0^{\infty}(\lim_{K\to\infty} Q_K(\theta))d\theta$$

$$= \int_0^{\infty} \lim\, [g(\theta)\int_0^{t_o} f_K(t|\theta)dt]\, d\theta$$

$$= \int_0^{\infty} g(\theta)\, [\lim \int_0^{t_o} f_K(t|\theta)dt]\, d\theta$$

$$= \int_0^{\infty} g(\theta) \cdot \begin{Bmatrix} 1 & \text{if } \theta < t_o \\ 0 & \text{if } \theta < t_o \end{Bmatrix} d\theta$$

$$= \int_0^{t_o} g(\theta) d\theta = G(t_o).$$

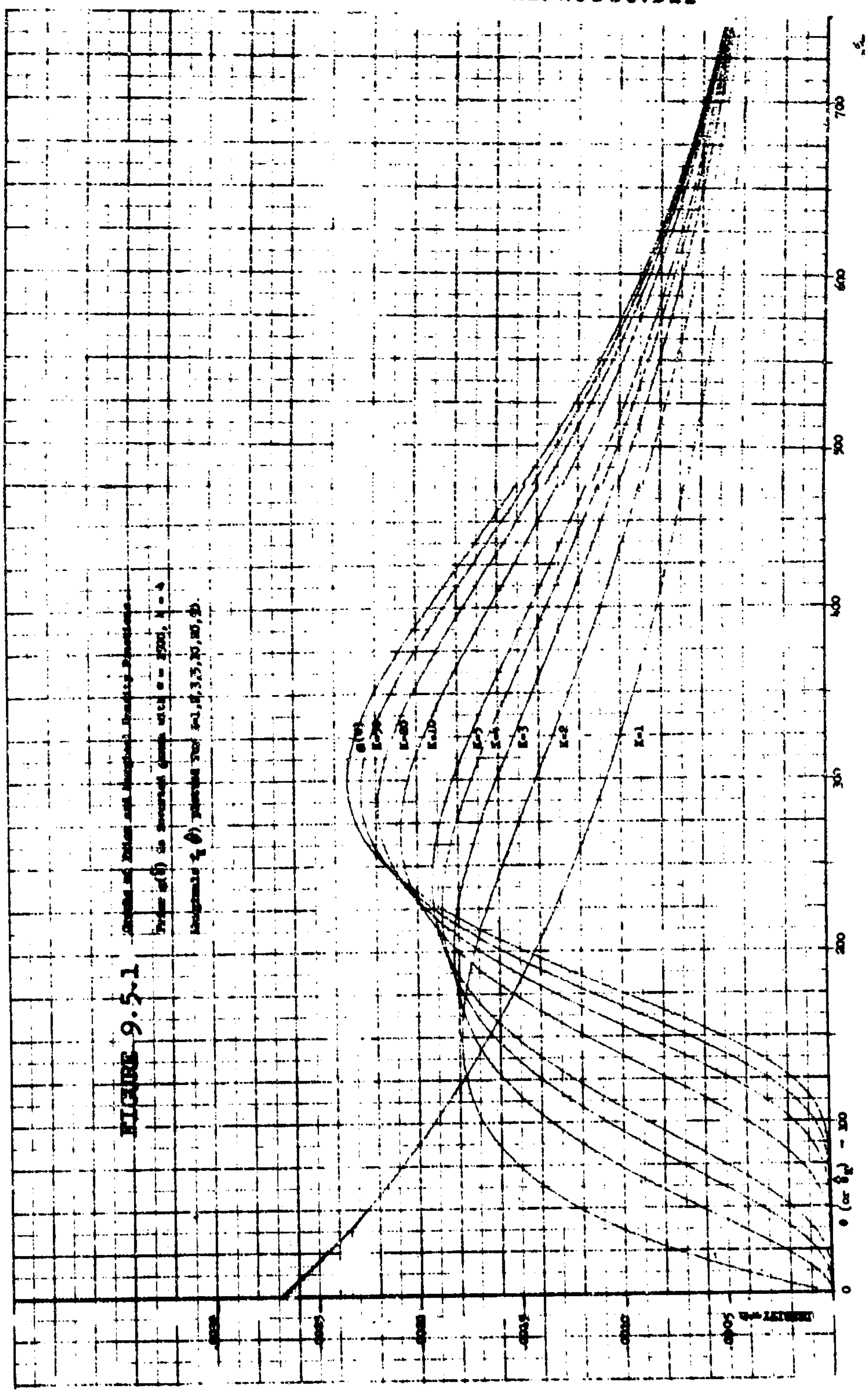


FIGURE 9.5.1

NOT REPRODUCIBLE

SECTION 10.0 COMPUTER PROGRAMS DEVELOPED

10.1 DATA ANALYSIS PROGRAMS

10.1.1 "KR"

This program uses data in pairs consisting of $\hat{\theta}$'s and K_i's to calculate the moments for the K-R method discussed in Sections 4.1 and 4.3.1 of this report. The program is designed such that if either the second or the fourth moment is negative, it terminates and prints "method not applicable." If both of these moments are positive, the sample $\hat{\beta}_1, \sqrt{\hat{\beta}_1}, \hat{\beta}_2, \hat{\kappa}$, and the estimated inverted gamma parameters, $\hat{\alpha}$ and $\hat{\lambda}$, are calculated. A list of the program follows.

```
IG          8:38    MON.    09/15/69

10 READ Y
20 PRINT"DATA SET NO.="Y
30 PRINT
40 LET M1=0
50 LET M2=0
60 LET M3=0
70 LET M4=0
80 READ K
90 PRINT" "," MTBF","NO. OF FAILURES"
100 PRINT"----------------------------"
110 FOR J=1 TO K
120 READ T,N
130 PRINT J,T,N
140 LET M1=M1+T
150 LET M2=M2+(T↑2)*N/(N+1)
160 LET M3=M3+(T↑3)*(N↑2)/((N+1)*(N+2))
170 LET M4=M4+(T↑4)*(N↑3)/((N+1)*(N+2)*(N+3))
180 NEXT J
190 PRINT"------------------------------"
200 PRINT
210 LET M1=M1/K
220 LET M2=M2/K
230 LET M3=M3/K
240 LET M4=M4/K
250 PRINT"M1="M1
260 PRINT"M2="M2
270 PRINT"M3="M3
280 PRINT"M4="M4
290 LET V2=M2-M1↑2
300 LET V3=M3-3*M1*M2+2*(M1↑3)
310 LET  V4=M4-4*M1*M3+6*(M1↑2)*M2-3*(M1↑4)
320 PRINT
330 PRINT"SAMPLE MOMENTS ABOUT MEAN"
340 PRINT
350 PRINT"SECOND="V2
360 PRINT"THIRD="V3
370 PRINT"FOURTH="V4
380 PRINT
390 IF V2<0 THEN 420
400 IF V4<0 THEN 420
410 GO TO 440
420 PRINT"METHOD NOT APPLICABLE"
430 GO TO 530
440 LET B1=(V3↑2)/(V2↑3)
450 LET S1=B1↑.5
460 LET B2=V4/(V2↑2)
470 PRINT"SAMPLE BETA1="B1
480 PRINT"SAMPLE SQ ROOT OF BETA1="S1
490 PRINT"SAMPLE BETA2="B2
500 LET K1=B1*((B2+3)↑2)
510 LET K2=4*(4*B2-3*B1)*(2*B2-3*B1-6)
520 PRINT"SAMPLE KAPPA="K1/K2
530 PRINT
540 PRINT"------------------------"
550 PRINT
560 PRINT"ESTIMATED INV GAMMA PARAMETERS"
570 PRINT
580 LET L=(2*M2-(M1↑2))/(M2-(M1↑2))
590 LET A=M1*(L-1)
600 PRINT"ALPHA="A
610 PRINT"LAMBDA="L
620 PRINT
630 PRINT"================================"
640 PRINT
650 GO TO 10
660 END
```

10.1.2 "BAYES"

This program first calculates $\hat{\alpha}$ and $\hat{\lambda}$ using equations 4.2.2.1.8 and 4.2.2.1.9. It then uses these estimates to calculate each of the $f(x_i)$'s according to equation 4.2.2.1.2. The χ^2 cells and the upper class limits are read into the program and it computes the χ^2 value by equation 4.2.3.1.1 and prints out a χ^2 value to be compared according to the method described in Section 4.2.3.1. Following is a list of the program and the printed results used in this study.

```
BAYES      8:31   WED.   10/08/69

10 DIM Z(50),E(50)
20 DIM D(20),O(20),N(20)
30 LET T=4320
40 READ A
50 PRINT "DATA SET NO."A
60 PRINT
70 LET Z0=0
80 FOR I=1 TO 50
90 LET Z(I)=0
100 NEXT I
110 LET M=0
120 LET S=0
130 LET Y=0
140 READ K,F
150 IF K=999 THEN 240
160 IF K>0 THEN 190
170 LET Z0=F
180 GO TO 200
190 LET Z(K)=F
200 LET Y=Y+F
210 LET M=M+K*F
220 LET S=S+(K↑2)*F
230 GO TO 140
240 LET M=M/Y
250 LET S=S/Y
260 LET V=S-(M↑2)
270 LET A=(M*T)/(V-M)
280 LET L=(A*M)/T
290 PRINT "ALPHA="A
300 PRINT "LAMBDA="L
310 PRINT
320 LET P=T/(T+A)
330 LET Q=A/(T+A)
340 LET E0=(Q↑L)
350 LET E0=E0*Y
360 FOR X=1 TO 50
370 LET F=(Q↑L)*(P↑X)
380 LET P1=1
390 LET P2=1
400 LET W=L+X-1
410 FOR I=1 TO X
420 LET P1=P1*W
430 LET P2=I*P2
440 LET W=W-1
450 NEXT I
460 LET E(X)=F*P1/P2
470 LET E(X)=E(X)*Y
480 NEXT X
490 PRINT
500 PRINT
510 READ C1
520 PRINT "NO. OF CHI SQUARE CELLS="C1
530 PRINT
540 PRINT "CELL NUMBER","UPPER LIMIT"
550 PRINT "--------------------------"
560 FOR I=1 TO C1-1
570 READ D(I)
580 PRINT I,D(I)
590 NEXT I
600 PRINT C1
610 PRINT
620 LET O(1)=Z0
630 LET N(1)=E0
```

```
640 FOR I=2 TO C1
650 LET O(I)=0
660 LET H(I)=0
670 NEXT I
680 FOR J=1 TO D(1)
690 LET O(1)=O(1)+Z(J)
700 LET H(1)=H(1)+E(J)
710 NEXT J
720 FOR I=2 TO C1-1
730 FOR J=D(I-1)+1 TO D(I)
740 LET O(I)=O(I)+Z(J)
750 LET H(I)=H(I)+E(J)
760 NEXT J
770 NEXT I
780 LET S1=0
790 LET T1=0
800 FOR I=1 TO C1-1
810 LET S1=S1+O(I)
820 LET T1=T1+H(I)
830 NEXT I
840 LET O(C1)=Y-S1
850 LET H(C1)=Y-T1
860 LET C2=0
870 PRINT "OBSERVED","EXPECTED"
880 PRINT "--------------------------"
890 FOR I=1 TO C1
900 PRINT O(I),H(I)
910 LET C2=C2+((O(I)-H(I))↑2)/H(I)
920 NEXT I
930 PRINT
940 PRINT "CHI SQUARE="C2
950 PRINT
960 PRINT"============================================"
970 PRINT
980 PRINT
990 GO TO 40
1000 DATA 1
1002 DATA 1,13,2,8,3,4,4,4,5,1,6,2,7,1
1004 DATA 8,4,11,1,13,1,17,1,19,1,999,999
1006 DATA 4,1,3,7,2
1008 DATA 1,11,2,11,3,10,4,7,5,5,6,4
1010 DATA 7,1,8,4,9,7,10,3,13,1,14,3
1012 DATA 15,1,16,3,19,1,22,1,44,1
1014 DATA 999,999
1016 DATA 8,1,2,3,4,6,12,15,3
1018 DATA 1,14,2,12,3,13,4,6,5,3,6,2
1020 DATA 7,1,8,1,11,2,13,2,17,1,28,1,29,1
1022 DATA 999,999
1024 DATA 6,1,2,3,4,8,4
1026 DATA 1,18,2,11,3,6,4,6,5,1,6,2,7,3
1028 DATA 8,1,9,1,10,1,12,1,999,999
1030 DATA 4,1,2,4,5
1032 DATA 1,22,2,7,3,1,4,4,5,2,6,1,7,3
1034 DATA 999,999
1036 DATA 4,1,2,4,6
1038 DATA 1,21,2,8,3,8,4,2,5,5,6,4,7,3
1040 DATA 8,2,9,1,19,1,999,999
1042 DATA 5,1,2,4,6,7
1044 DATA 1,18,2,10,3,5,4,2,5,3,6,3,7,3,9,4
1046 DATA 10,4,11,1,12,3,13,1,15,1,999,999
1048 DATA 5,1,3,7,10,8
1050 DATA 1,16,2,10,3,7,4,5,5,3,6,3,7,1
1052 DATA 8,1,9,1,12,1,15,1,33,1,999,999
1054 DATA 4,1,3,5
9999 END
```

```
BAYES      8:37   WED.   10/08/69

DATA SET NO. 1

ALPHA= 1279.67
LAMBDA= [illegible].24991


NO. OF CHI SQUARE CELLS= 4

CELL NUMBER     UPPER LIMIT
-------------------------
 1               1
 2               3
 3               7
 4

OBSERVED        EXPECTED
-------------------------
 13              12.7267
 12              9.95358
 8               11.0720
 8               7.24698

CHI SQUARE= 1.35756

===============================================

DATA SET NO. 2

ALPHA= 758.591
LAMBDA= 1.12716


NO. OF CHI SQUARE CELLS= 8

CELL NUMBER     UPPER LIMIT
-------------------------
 1               1
 2               2
 3               3
 4               4
 5               6
 6               12
 7               15
 8

OBSERVED        EXPECTED
-------------------------
 11              17.0014
 11              7.52091
 10              6.67370
 7               5.03914
 9               9.55016
 15              16.4530
 5               8.07046
 6               6.06035

CHI SQUARE= 6.07651

===============================================
```

DATA SET NO. 3

ALPHA= 727.436
LAMBDA= .756319

NO. OF CHI SQUARE CELLS= 6

CELL NUMBER	UPPER LIMIT
1	1
2	2
3	3
4	4
5	8
6	

OBSERVED	EXPECTED
14	22.4572
12	6.63259
13	5.2156
6	4.19199
7	10.4516
7	10.051

CHI SQUARE= 21.9927

==

DATA SET NO. 4

ALPHA= 3609.57
LAMBDA= 2.60495

NO. OF CHI SQUARE CELLS= 4

CELL NUMBER	UPPER LIMIT
1	1
2	2
3	4
4	

OBSERVED	EXPECTED
18	15.8811
11	9.14851
12	13.4907
10	12.4797

CHI SQUARE= 1.31484

==

DATA SET NO. 5

ALPHA= 7584.73
LAMBDA= 4.03817

NO. OF CHI SQUARE CELLS= 4

CELL NUMBER	UPPER LIMIT
1	1
2	2
3	4
4	

OBSERVED	EXPECTED
22	15.9718
7	8.67811
5	10.3853
6	4.96478

CHI SQUARE= 5.60818

==

DATA SET NO. 6

ALPHA= 2251.61
LAMBDA= 1.73419

NO. OF CHI SQUARE CELLS= 5

CELL NUMBER	UPPER LIMIT
1	1
2	2
3	4
4	6
5	

OBSERVED	EXPECTED
21	18.3683
8	8.79374
10	12.7938
9	7.3346
7	7.70951

CHI SQUARE= 1.50224

==

DATA SET NO. 7

ALPHA= 1767.69
LAMBDA= 1.[illegible]7662

NO. OF CHI SQUARE CELLS= 5

CELL NUMBER	UPPER LIMIT
1	1
2	3
3	7
4	10
5	

OBSERVED	EXPECTED
18	13.2822
15	14.8425
11	18.7591
8	6.23432
6	4.88186

CHI SQUARE= 5.64293

==

DATA SET NO. 8

ALPHA= 770.804
LAMBDA= .692296

NO. OF CHI SQUARE CELLS= 4

CELL NUMBER	UPPER LIMIT
1	1
2	3
3	5
4	

OBSERVED	EXPECTED
16	21.4834
17	10.056
8	6.11726
9	12.3434

CHI SQUARE= 7.67972

==

10.2 SIMULATION

10.2.1 "BAYES 2"

This is the simulation described in Section 5.2.2.1. One form of the program, as it was used, follows.

```
      DIMENSION N(6),K(4),NCEL(6)
      DIMENSION IPF(6,4,2),THET(200)
      DIMENSION GMARK(19),CS1(6),CS2(6),Q(200)
      DIMENSION ZMARK(19),NGTAL(20),EG(20),EN(20),EW(20),DN(19),DW(19)
      DIMENSION INPF(6,4,2),IWPF(6,4,2),ICOLUM(4)
      DIMENSION IC1(6),IC2(6),IC3(6)
      DIMENSION XGAM(99),XWEIB(99),XNRM(99)
      DIMENSION YGAM(99,4),YWEIB(99,4),YNRM(99,4)
      DIMENSION IUPF(6,4,2),IC4(6)
      DIMENSION EU(20),DU(20)
      IX=48828125
    5 FORMAT(6I5)
    7 FORMAT(6I5)
   10 FORMAT(4I5)
   15 FORMAT(6F10.3)
   20 FORMAT(10F8.5/10F8.5/10F8.5/10F8.5/10F8.5)
   25 FORMAT (F10.2,I10,4F10.5)
   30 FORMAT (1H1/15H INVERTED GAMMA,20X,7H ALPHA=,3X,F10.5,10X,6H LMDA=
     1,4X,I10  /7H LOGNRM,28X,4H MU=,6X,F10.5,10X,7H SIGMA=,3X,F10.5/6H
     2WEIBULL,22X,7H ALPHA=,3X,F15.5,10X,6H BETA=,F10.5///)
   35 FORMAT(10X,3H N=,I4//10X,3H K=,I4//)
   45 FORMAT(////10H CELL NO. ,10H   CLASS  ,10H OBSERVED ,10H  EXPECTED
     1,10H  EXPECTED,10H EXPECTED /10X,10H   MARK   ,10X,10H INV GAMMA,
     210H LOG NORML,10H WEIBULL  //)
   50 FORMAT(I10,F10.5,I10,3F10.5)
   55 FORMAT(I10,10X,I10,3F10.5////20X,14H CHI SQ VALUES//20X,10H INV GA
     1MMA,10H LOG NORML,10H  WEIBULL /20H COMPUTED CHI SQUARE,3F10.5/
     220H 90 PERCENT LEVEL    ,3F10.5/20H 95 PERCENT LEVEL    ,3F10.5)
   60 FORMAT(////30H ******************************/////)
   70 FORMAT (///20X,25H SUMMARY OF RESULTS AFTER,I5,8H  TRIALS/
     120X,25H LEVEL 1=.90  LEVEL 2=.95/    20X, 7H ALPHA=,F10.5,5X,8H LA
     2MBDA=,I10///)
   75 FORMAT(20X,7H LEVEL=,I2///20X,15H INVERTED GAMMA///)
   77 FORMAT (10H         N=,6I10//10H         K=/)
   80 FORMAT (7I10/)
   85 FORMAT(///20X,11H LOG NORMAL///)
   87 FORMAT(///20X,8H WEIBULL///)
   88 FORMAT (///20X,8H UNIFORM///)
   90 FORMAT(///20X,10H **********////)
   91 FORMAT (7I10/8X,2H L,6I10/8X,2H W,6I10/8X,2H U,6I10//)
   96 FORMAT (///20X,12H ALL 4 TESTS///)
   97 FORMAT (I5,F8.2,8X,F8.2,8X,F8.2)
   99 FORMAT (34H TABLE OF SAMPLE PERCENTAGE POINTS/
     123H OF THE MARGINAL FOR K=,I5//5X,16H INV GAMMA       ,16H LOG NORM
     2AL       ,8H WEIBULL//)
  142 FORMAT (10F8.2/10F8.2/10F8.2/10F8.2/10F8.2/10F8.2/10F8.2/10F8.2/
     C10F8.2/9F8.2)
      READ (5,5) (N(J),J=1,6)
```

```
      READ (5,7) (NCEL(J),J=1,6)
      READ (5,10) (K(J),J=1,4)
      READ (5,15) (CS1(J),J=1,6)
      READ (5,15) (CS2(J),J=1,6)
  100 READ (5,25) ALP,LAM,AMU,SIG,WALP,WBET
      IF(ALP.EQ.999.) GO TO 999
      DO 102 J=1,4
      READ (5,142) (XGAM(I),I=1,99)
      READ (5,142) (XNRM(I),I=1,99)
      READ (5,142) (XWEIB(I),I=1,99)
      KO=K(J)
      WRITE (6,99) KO
      DO 98 I=1,99
      WRITE (6,97) I,XGAM(I),XNRM(I),XWEIB(I)
   98 CONTINUE
      WRITE (6,60)
      DO 103 I=1,99
      XG=XGAM(I)
      YGAM(I,J)=XG
      XN=XNRM(I)
      YNRM(I,J)=XN
      XW=XWEIB(I)
  103 YWEIB(I,J)=XW
  102 CONTINUE
      AUNIF=0.
      BUNIF=2.*((3./2.)**.5)*500.
      WALP=(AMU/GAMMA(1.+1./WBET))**WBET
      WRITE (6,30) ALP,LAM,AMU,SIG,WALP,WBET
      AMU1=ALOG(AMU**2/((AMU**2+SIG**2)**.5))
      SIG1=(ALOG((SIG**2+AMU**2)/AMU**2))**.5
      AMU=AMU1
      SIG=SIG1
      DO 710 I=1,6
      DO 720 J=1,4
      DO 730 KK=1,2
      IPF(I,J,KK)=0
      INPF(I,J,KK)=0
      IWPF(I,J,KK)=0
      IUPF(I[illegible]K )=0
  730 CONTINUE
  720 CONTINUE
  710 CONTINUE
      ITEN=0
      DO 700 LOOP=1,1000
      I=6
      NCELL=NCEL(I)
  117 NO=N(I)
      ZN=NCELL
```

```
      MCELL=NCELL-1
      XNO=NO
      DO 120 J=1,4
      KO=K(J)
      XKO=KO
      DO 127 IA=1,99
      XG=YGAM(IA,J)
      XN=YNRM(IA,J)
      XW=YWEIB(IA,J)
      XGAM(IA)=XG
      XNRM(IA)=XN
      XWEIB(IA)=XW
  127 CONTINUE
      DO 119 JX=1,MCELL
      ZJ=JX
      ZMARK(JX)=ZJ/ZN
      ZM=ZMARK(JX)
      DO 123 L=1,99
      XL=L
      XL=XL/100.
      IF (ZM.GE.XL) GO TO 123
      GM=XGAM(L-1)+((ZM-XL+.01)/.01)*(XGAM(L)-XGAM(L-1))
      GO TO 118
  123 CONTINUE
  118 GMARK(JX)=GM
  119 CONTINUE
      DO 121 JX=1,MCELL
      U=GMARK(JX)
      CALL LOOKUP (XWEIB,U,DWE)
      DW(JX)=DWE
  121 CONTINUE
      DO 122 JX=1,MCELL
      U=GMARK(JX)
      CALL LOOKUP (XNRM,U,DNO)
      DN(JX)=DNO
  122 CONTINUE
      DO 740 JX=1,MCELL
      U=GMARK(JX)
      IF (U.LT.BUNIF) GO TO 750
      DU(JX)=1.
      GO TO 740
  750 IF (U.GT.AUNIF) GO TO 760
      DU(JX)=0.
      GO TO 740
  760 DU(JX)=(U-AUNIF)/(BUNIF-AUNIF)
  740 CONTINUE
      EG(1)=XNO*ZMARK(1)
      EG(NCELL)=XNO*(1.-ZMARK(MCELL))
```

```
      EN(1)=XNO*DN(1)
      EN(NCELL)=XNO*(1.-DN(MCELL))
      EW(1)=XNO*DW(1)
      EW(NCELL)=XNO*(1.-DW(MCELL))
      EU(1)=XNO*DU(1)
      EU(NCELL)=XNO*(1.-DU(MCELL))
      IF (NCELL.EQ.2) GO TO 126
      DO 124 JX=2,MCELL
      EG(JX)=XNO*(ZMARK(JX)-ZMARK(JX-1))
      EW(JX)=XNO*(DW(JX)-DW(JX-1))
      EN(JX)=XNO*(DN(JX)-DN(JX-1))
      EU(JX)=XNO*(DU(JX)-DU(JX-1))
  124 CONTINUE
  126 DO 125 II=1,NC
      Q(II)=0.
  125 CONTINUE
      DO 130 II=1,NO
      CALL RANDO(IX,IY,Y)
      IX=IY
      CALL VRTGAM (ALP,LAM,Y,THETA)
      THET(II)=THETA
      DO 150 JJ=1,KO
      CALL RANDO(IX,IY,Z)
      IX=IY
      S=-THETA*ALOG(Z)
      Q(II)=Q(II)+S
  150 CONTINUE
      Q(II)=Q(II)/XKO
  130 CONTINUE
      DO 170 IT=1,NCELL
      NGTAL(IT)=0
  170 CONTINUE
      DO 180 II=1,NC
      DO 190 IT=1,MCELL
      IF (Q(II).LT.GMARK(IT)) GO TO 200
      GO TO 190
  200 NGTAL(IT)=NGTAL(IT)+1
      GO TO 180
  190 CONTINUE
      NGTAL(NCELL)=NGTAL(NCELL)+1
  180 CONTINUE
      CSIG=0.
      CSN=0.
      CSW=0.
      CSU=0.
      DO 230 IT=1,NCELL
      F=NGTAL(IT)
      CSIG=CSIG+((F-EG(IT))**2)/(EG(IT))
```

```
      CSN=CSN+((F-EN(IT))**2)/(EN(IT))
      CSW=CSW+((F-EW(IT))**2)/(EW(IT))
  230 CONTINUE
      NGTAL(19)=NGTAL(19)+NGTAL(20)
      EU(19)=EU(19)+EU(20)
  231 DO 232 IT=1,19
      F=NGTAL(IT)
      CSU=CSU+((F-EU(IT))**2)/(EU(IT))
  232 CONTINUE
      IF (CSIG.GT.CS1(I)) GO TO 260
      IPF(I,J,1)=IPF(I,J,1)+1
  260 IF (CSIG.GT.CS2(I)) GO TO 280
      IPF(I,J,2)=IPF(I,J,2)+1
  280 IF (CSN.GT.CS1(I)) GO TO 300
      INPF(I,J,1)=INPF(I,J,1)+1
  300 IF (CSN.GT.CS2(I)) GO TO 320
      INPF(I,J,2)=INPF(I,J,2)+1
  320 IF (CSW.GT.CS1(I)) GO TO 340
      IWPF(I,J,1)=IWPF(I,J,1)+1
  340 IF (CSW.GT.CS2(I)) GO TO 765
      IWPF(I,J,2)=IWPF(I,J,2)+1
  765 IF (CSU.GT.25.989) GO TO 767
      IUPF(I,J,1)=IUPF(I,J,1)+1
  767 IF (CSU.GT.28.869) GO TO 120
      IUPF(I,J,2)=IUPF(I,J,2)+1
  120 CONTINUE
      ITEN=ITEN+1
      IF (ITEN.LT.50) GO TO 700
      WRITE (6,70) LCOP,ALP,LAM
      DO 370 KK=1,2
      WRITE (6,75) KK
      WRITE (6,77) (N(M),M=1,6)
      DO 380 J=1,4
      DO 390 I=1,6
      ICOLUM(I)=IPF(I,J,KK)
  390 CONTINUE
      WRITE (6,80) K(J),(ICOLUM(I),I=1,6)
  380 CONTINUE
      WRITE(6,85)
      WRITE (6,77) (N(M),M=1,6)
      DO 480 J=1,4
      DO 490 I=1,6
      ICOLUM(I)=INPF(I,J,KK)
  490 CONTINUE
      WRITE (6,80) K(J),(ICOLUM(I),I=1,6)
  480 CONTINUE
      WRITE(6,87)
      WRITE (6,77) (N(M),M=1,6)
```

```
      DO 580 J=1,4
      DO 590 I=1,6
      ICOLUM(I)=IWPF(I,J,KK)
  590 CONTINUE
      WRITE (6,80) K(J),(ICCLUM(I),I=1,6)
  580 CONTINUE
      WRITE (6,88)
      WRITE (6,77) (N(M),M=1,6)
      DO 880 J=1,4
      DO 890 I=1,6
      ICCLUM(I)=IUPF(I,J,KK)
  890 CONTINUE
      WRITE (6,80) K(J),(ICCLUM(I),I=1,6)
  880 CONTINUE
      WRITE (6,96)
      WRITE (6,77) (N(M),M=1,6)
      DO 680 J=1,4
      DO 690 I=1,6
      IC1(I)=IPF(I,J,KK)
      IC2(I)=INPF(I,J,KK)
      IC3(I)=IWPF(I,J,KK)
      IC4(I)=IUPF(I,J,KK)
  690 CONTINUE
      WRITE (6,91) K(J),(IC1(L),L=1,6),(IC2(L),L=1,6),(IC3(L),L=1,6)
     C,(IC4(L),L=1,6)
  680 CONTINUE
      WRITE(6,90)
  370 CONTINUE
      WRITE (6,692) IX
  692 FORMAT (10X,4H IX=,I20)
      WRITE (6,60)
      ITEN=0
  700 CONTINUE
      GO TO 100
  999 STOP
      END
```

```
      SUBROUTINE LOOKUP (ARRAY,U,P)
      DIMENSION ARRAY(99)
      DO 100 L=1,99
      IF (U.GE.ARRAY(L)) GO TO 100
      P1=L-1
      P1=P1/100.
      P=P1+.01*(U-ARRAY(L-1))/(ARRAY(L)-ARRAY(L-1))
      GO TO 200
  100 CONTINUE
  200 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE RANDO (IX,IY,YFL)
      IY=IX*65539
      IF(IY) 5,6,6
    5 IY=IY+2147483647+1
    6 YFL=IY
      YFL=YFL*.4656613E-9
      RETURN
      END
```

```
      SUBROUTINE VRTGAM(ALP,LAM,P,THETA)
      ALAM=LAM
      AMU=ALP/(ALAM-1.)
   50 X1=1./AMU
      M=LAM-1
      ICOUNT=0
  100 V=1.
      S=1.
      DO 150 K=1,M
      FK=K
      V=V*(ALP*X1)/FK
      S=S+V
  150 CONTINUE
      X2=X1+((S*EXP(-ALP*X1))-(1.-P))/(ALP*V*EXP(-ALP*X1))
      X3=X2-X1
      FABS=ABS(X3)
      IF (FABS.LT.0.000001) GO TO 200
      X1=X2
      ICOUNT=ICOUNT+1
      IF (ICOUNT.EQ.20) GO TO 250
      GO TO 100
  200 T=1./X2
      IF (T.GT.0.) GO TO 300
  250 WRITE(6,240) P
  240 FORMAT (//12H HUNG UP ON ,F10.5//)
  260 P=RAND(X)
      GO TO 50
  300 THETA=T
      RETURN
      END
```

10.2.2 "BAYES 12"

This is the simulation program described in Section 5.2.2.2. A list of the program follows.

```
      DIMENSION VRTG(51),WEIB(51),ZLOG(51)
      DIMENSION IRAND(51),XVRTG(51),XWEIB(51),XZLOG(51)
      DIMENSION ID(20),IG(20),MG(20),HW(20),HL(20)
      DIMENSION CUMIG(51)
      IX=48828125
  100 READ (5,25) ALP,ALAM,AMU,SIG,WALP,WBET,TIME
   25 FORMAT (7F10.5)
      IF (ALP.EQ.999.) GO TO 999
      WRITE (6,94)
   94 FORMAT (///5X,20H ********************////)
      WALP=(AMU/GAMMA(1.+1./WBET))**WBET
      WRITE (6,30) ALP,ALAM,AMU,SIG,WALP,WBET
   30 FORMAT (10X,2F10.5//10X,2F10.5//10X,2F10.5///)
      AMU1=ALOG(AMU**2/((AMU**2+SIG**2)**.5))
      SIG1=(ALOG((SIG**2+AMU**2)/AMU**2))**.5
      AML=AMU1
      SIG=SIG1
      DO 110 I=1,8
      VRTG(I)=0.
      WEIB(I)=0.
  110 ZLOG(I)=0.
      ZG=0.
      ZW=0.
      ZL=0.
C
C         MARGINAL FOR INV GAMMA PRIOR
C
      P=TIME/(TIME+ALP)
      Q=ALP/(TIME+ALP)
      VRTG(1)=Q**ALAM
      XTAL=VRTG(1)
      CUMIG(1)=XTAL
      DO 120 J=1,7
      F=(Q**ALAM)*(P**J)
      P1=1.
      P2=1.
      XJ=J
      W=ALAM+XJ-1.
      DO 130 I=1,J
      P1=P1*W
      XI=I
      P2=XI*P2
      W=W-1.
  130 CONTINUE
      VRTG(J+1)=F*P1/P2
      XTAL=XTAL+VRTG(J+1)
      CUMIG(J+1)=XTAL
  120 CONTINUE
```

```
      ZG=1.-XTAL
C
C        MARGINAL FOR WEIBULL PRIOR
C
      DO 200 I=1,10000
      CALL RANDO(IX,IY,P)
      IX=IY
      THETA=(WALP*ALOG(1./P))**(1./WBET)
      K=0
      CLOCK=0.
  250 CALL RANDO(IX,IY,F)
      IX=IY
      T=-THETA*ALOG(F)
      CLOCK=CLOCK+T
      IF(CLOCK.GT.TIME) GO TO 300
      IF(K.LT.8) GO TO 201
      ZW=ZW+1.
      GO TO 200
  201 K=K+1
      GO TO 250
  300 WEIB(K+1)=WEIB(K+1)+1.
  200 CONTINUE
      DO 310 I=1,8
  310 WEIB(I)=WEIB(I)/10000.
      ZW=ZW/10000.
C
C        MARGINAL FOR LOGNORMAL PRIOR
C
      DO 400 I=1,10000
      A=0.
      DO 450 J=1,12
      CALL RANDO(IX,IY,Y)
      IX=IY
  450 A=A+Y
      THETA=EXP((A-6.0)*SIG+AMU)
      K=0
      CLOCK=0.
  460 CALL RANDO(IX,IY,F)
      IX=IY
      T=-THETA*ALOG(F)
      CLOCK=CLOCK+T
      IF(CLOCK.GT.TIME) GO TO 480
      IF (K.LT.8) GO TO 470
      ZL=ZL+1.
      GO TO 400
  470 K=K+1
      GO TO 460
  480 ZLOG(K+1)=ZLOG(K+1)+1.
```

```
  400 CONTINUE
      DO 410 I=1,8
  410 ZLOG(I)=ZLOG(I)/10000.
      ZL=ZL/10000.
      DO 500 L=1,6
      READ(5,5) NO
    5 FORMAT(I5)
      WRITE (6,96) NO
   96 FORMAT (//10X,3H N=,I5)
      READ(5,10) IC1,CS1,CS2
   10 FORMAT(I10,2F10.5)
      JC1=IC1-1
      DO 600 I=1,JC1
      READ(5,5) ID(I)
      ID(I)=ID(I)+1
  600 CONTINUE
      XNO=NO
      DO 700 I=1,8
      XVPTG(I)=VPTG(I)*XNO
      XWEIR(I)=WEIR(I)*XNO
      XZLOG(I)=ZLOG(I)*XNO
  700 CONTINUE
      XZG=ZG*XNO
      XZW=ZW*XNO
      XZL=ZL*XNO
      IPG1=0
      IPW1=0
      IPL1=0
      IPG2=0
      IPW2=0
      IPL2=0
      DO 900 LOOP=1,1000
      DO 490 I=1,8
  490 IRAND(I)=0
      IR=0
      DO 610 I=1,IC1
      IG(I)=0
      HG(I)=0.
      HW(I)=0.
      HL(I)=0.
  610 CONTINUE
      DO 510 II=1,NO
      CALL RANDO(IX,IY,Z)
      IX=IY
      DO 520 J=1,8
      IF (Z.LT.CUMIG(J)) GO TO 530
  520 CONTINUE
      IR=IR+1
```

```
      GO TO 510
  530 IRAND(J)=IRAND(J)+1
  510 CONTINUE
      IF (LOOP.NE.1) GO TO 910
      WRITE(6,99)
   99 FORMAT (///18X,2H K,5X,9H EXPECTED,6X,9H EXPECTED,6X,9H EXPECTED,
     16X,9H OBSERVED/25X,10H INV GAMMA,5X,10H  WEIBULL ,5X,10H LOG NORML
     2//)
      DO 800 I=1,8
      IJ=I-1
      WRITE (6,98) IJ,XVRTG(I),XWEIB(I),XZLOG(I),IRAND(I)
   98 FORMAT( 10X,I10,5X,F10.5,5X,F10.5,5X,F10.5,5X,I10)
  800 CONTINUE
      WRITE (6,97) XZG,XZW,XZL,IR
   97 FORMAT( 10X,8H OVER  7,7X,F10.5,5X,F10.5,5X,F10.5,5X,I10)
  910 ID1=ID(1)
      DO 620 J=1,ID1
      IG(1)=IG(1)+IRAND(J)
      HG(1)=HG(1)+XVRTG(J)
      HW(1)=HW(1)+XWEIB(J)
      HL(1)=HL(1)+XZLOG(J)
  620 CONTINUE
      IGS=IG(1)
      HGS=HG(1)
      HWS=HW(1)
      HLS=HL(1)
      IF (IC1.EQ.2) GO TO 645
      DO 630 I=2,IC1
      J1=ID(I-1)+1
      J2=ID(I)
      DO 640 J=J1,J2
      IG(I)=IG(I)+IRAND(J)
      HG(I)=HG(I)+XVRTG(J)
      HW(I)=HW(I)+XWEIB(J)
      HL(I)=HL(I)+XZLOG(J)
  640 CONTINUE
      IGS=IGS+IG(I)
      HGS=HGS+HG(I)
      HWS=HWS+HW(I)
      HLS=HLS+HL(I)
  630 CONTINUE
  645 IG(IC1)=NO-IGS
      HG(IC1)=XNO-HGS
      HW(IC1)=XNO-HWS
      HL(IC1)=XNO-HLS
      CSG=0.
      CSW=0.
      CSL=0.
```

```
      DO 650 I=1,IC1
      G=IG(I)
      CSG=CSG+((G-HG(I))**2)/HG(I)
      CSW=CSW+((G-HW(I))**2)/HW(I)
      CSL=CSL+((G-HL(I))**2)/HL(I)
  650 CONTINUE
      IF (IJOP.NE.1) GO TO 920
      WRITE(6,93)CSG,CSW,CSL,CS1,CS1,CS1,CS2,CS2,CS2
   93 FORMAT(///10X,18H CHI  SQUARE VALUES//10X,3F10.5/10X,3F10.5/10X,3F1
     C0.5//)
      WRITE (6,95)
   95 FORMAT (///10X,10H **********////)
  920 IF (CSG.GT.CS1) GO TO 930
      IPG1=IPG1+1
  930 IF (CSG.GT.CS2) GO TO 940
      IPG2=IPG2+1
  940 IF (CSW.GT.CS1) GO TO 950
      IPW1=IPW1+1
  950 IF (CSW.GT.CS2) GO TO 960
      IPW2=IPW2+1
  960 IF (CSL.GT.CS1) GO TO 970
      IPL1=IPL1+1
  970 IF (CSL.GT.CS2) GO TO 900
      IPL2=IPL2+1
  900 CONTINUE
      WRITE (6,92) IPG1,IPG2,IPW1,IPW2,IPL1,IPL2
   92 FORMAT (///10X,8H SUMMARY///10X,10H INV GAMMA,5X,2I5//
     C10X,10H WEIBULL   ,5X,2I5//10X,10H LOG NORML,5X,2I5//)
      WRITE (6,94)
  500 CONTINUE
      GO TO 100
  999 STOP
      END
```

```
      SUBROUTINE RANDO (IX,IY,YFL)
      IY=IX*65539
      IF(IY) 5,6,6
    5 IY=IY+2147483647+1
    6 YFL=IY
      YFL=YFL*.4656613E-9
      RETURN
      END
```

SECTION 11.0 REFERENCES AND BIBLIOGRAPHY

Section 11.1 of this report contains a complete list of all the references used for this study, whereas, Section 11.2 is a general bibliography containing sources of material on Bayes reliability, in general, and is meant only for the readers use and convenience.

11.1 REFERENCES

1. Bhattacharya, S. K. Bayesian Approach to Life Testing and Reliability Estimation. Journal of the American Statistical Association, March 1967.

2. Elderton, W. P., Frequency Curves and Correlation, 4th edition, Harper Press, Washington, D.C., 1953.

3. Feller, William. An Introduction to Probability Theory and Its Applications, Vol. I, 3rd edition, John Wiley & Sons, Inc., New York, 1968.

4. Krutchkoff, R. G. and Rutherford, J. R. The Empirical Bayes Approach: Estimating the Prior Distribution, Biometrika, Vol. 54, 1967.

5. Rolph, John E. Bayesian Estimation of Mixing Distributions. The Annals of Mathematical Statistics, Vol. 39, No. 4, 1968.

6. Schafer, R. E. and Collins, J. A Note on the Moments of the Inverted Gamma Distribution. Unpublished.

7. Teicher, Henry. On the Mixture of Distributions. Annals of Mathematical Statistics, Vol. 31, 1960.

8. Teicher, Henry. Identifiability of Mixtures. The Annals of Mathematical Statistics, 1961.

9. Tucker, H. G. An Estimate of the Compounding Distribution of a Compound Poisson Distribution. Theory of Probability and Its Applications, Vol. 8, 1963.

10. Whittle, P. . On the Smoothing of Probability Density Functions. Journal of the Royal Statistical Society, Vol. 20, 1958.

11.2 BIBLIOGRAPHY

1. Babillis, R.A. and Smith, A.M. "Application of Bayesian Statistics in Reliability Measurement", Reliability and Maintainability, 4th Annual Conference, July 1965.

2. Balaban, H. "A Bayesian Approach to Reliability Demonstration", Annals of Assurance Sciences, Vol 8, 1969.

3. Balaban, H.S. "A Bayesian Approach for Designing Component Life Tests", IEEE 1967 Symposium on Reliability, 1967.

4. Balaban, H. "Bayesian Methods in Reliability and Life Testing", Electronics Division Newsletter, No. 2, American Society for Quality Control, October 1968, P. 3-8.

5. Bhattacharya, S.K. "Bayesian Approach to Life Testing and Reliability Estimation", Journal of the American Statistical Association, Vol 62, March 1967.

6. Bonis, A.J. "Bayesian Reliability Demonstration Plans", Annals of Reliability and Maintainability, Vol 5, 1965.

7. Breipohl, A.M. "A Consideration of the Bayesian Approach in Reliability Evaluation", IEEE Transactions on Reliability, Vol R-14, October 1965.

8. Breipohl, A.M., Prairie, R.R. and Zimmer, W.J. "A Consideration of the Bayesian Approach in Reliability Evaluation", IEEE, Transactions of Reliability, October 1965.

9. Breipohl, A.M. and McCormick, W.C., Jr. "Bayesian Estimation of Time-Varying Reliability", Annals of Assurance Sciences, Vol 1 No. 2, 1968.

10. Brender, D.M. "The Prediction and Measurement of System Availability: A Bayesian Treatment", IEEE Transactions, September 1968.

11. Brender, D.M. "The Bayesian Assessment of System Availability: Advanced Applications and Techniques", IEEE Transactions, September 1968.

12. Brender, D.M. "Reliability Testing in a Bayesian

Context", IEEE International Convention Record Vol 14, Part 9, 1966.

13. Bricker, J.L. and Schafer, R.E. "A Sequential Bayes Test of a Composite Hypothesis", Submitted for Publication, May 1969.

14. Briggs, W.G. "Reliability Statements Under Uncertainty", The Logistics Review, Vol 3, January 1967.

15. Collins, J.R. and Schafer, R.E. "A Note on the Moments of the Inverted Gamma Distribution", Unpublished.

16. Cornfield, J. "A Bayesian Test of Some Classical Hypothesis with Applications to Classical Sequential Clinical Trials", Journal of American Statistical Association, September 1966.

17. Cox, D.R. "Serial Sampling Acceptance Schemes Derived from Bayes Theorem", Technometrics, Vol 2, No. 3, August 1960.

18. Deely, J.J. and Zimmer, W.J. "Some Comparisons of Bayesian and Classical Confidence Intervals in the Exponential Cases", Annals of Assurance Sciences, Vol 1, No. 2, 1968.

19. Deely, J.J. and Zimmer, W.J. "Partial Prior Information and Shorter Confidence Intervals", Annals of Assurance Sciences, Vol 8, 1969.

20. De Hardt, J.H. and McLaughlin, H.D. "Using Bayesian Methods to Select a Design with Known Reliability Without a Confidence Coefficient", Annals of Reliability and Maintainability, Vol 5, July 1965.

21. Drake, A.W. "Bayesian Statistics for the Reliability Engineer", Proceedings of 1966 Symposium on Reliability, January 1966.

22. Earnest, C.M. "Estimating Reliability after Corrective Action: A Bayesian Viewpoint", Postgraduate Thesis U.S. Naval School, May 1966.

23. Elfving, G. "Robustness of Bayes Decisions Against the Choice of Prior", AD 642 613, November 1966.

24. Fox, B.L. "A Bayesian Approach to Reliability Assessment", Performed Under Contract NASR-21 (11), August 1966.

25. Girshick, M.A. and Rubin, H. "A Bayes Approach to a

Quality Control Model", Annals of Mathematical Statistics, Vol 23, 1952.

26. Greenberg, S.A. and Rosinoff, T. "Optimum Decision Criteria for Reliability Tests", Mitre Corporation, May 1967.

27. Guthrie, D. and Johns, M.W. "Bayes' Acceptance Sampling Procedures for Large Lots", Annals of Mathematical Statistics, 1959.

28. Hald, A. "Bayesian Single Sampling Attribute Plans for Discrete Prior Distribution", Prepared Under Contract NONR-N62558-3073, June 1964.

29. Hald, A. "Bayesian Single Sampling Attribute Plans for Continuous Prior Distributions", Technometrics, Vol 10, No. 4, November 1968.

30. Hald, A. "Asymptotic Properties of Bayesian Single Sampling Plans", Journ. Roy. Statist. Soc. B, 1967.

31. Hamilton, C.W. "Bayesian Procedures and Reliability Information", Aerospace Reliability and Maintainability Conference, May 1963.

32. Hamilton, C.W. and Drennan, J.E. "Research Toward a Bayesian Procedure for Calculating System Reliability", Proceeds Third Annual Aerospace Reliability and Maintainability Conference, 1964.

33. Harris, C. and Singpurwalla, N. "Life-Time Distributions Derived from Stochastics Hazard Functions", IEEE Transcript on Reliability, R17 No. 2, June 1968.

34. Hartley, H.O. "In Dr. Bayes' Consulting Room", The American Statistician, Vol 17, February 1963.

35. Hershman, R.L. and Froitag, M. "A Bayesian Model for Troubleshooting Electronic Equipment", Research Report January-July 1966, November 1966.

36. Hillborn, C.G. and Lainiotis, D.G. "Optimal Adaptive Filter Realizations for Sampled Stochastic Processes with an Unknown Parameter", Texas University Thesis, July 1967.

37. Holla, M.S. "Bayesian Estimates of the Reliability Function", Australian Journal of Statistics, Vol 8, 1966.

38. Howard, R.A. "Bayesian Decision Models for System Engineering", IEEE Transactions on Systems Science and Cybernetics, Vol SSC-1, November 1965.

39. Keefe, A.A. and Ohara, H.T. "Confidence Intervals for System Reliability from Component Testing Using Converse Hypergeometric Probability Distributions", Journal of the Electronics Division, ASQC, Vol 5, April 1967.

40. Larson, Harold J. "Bayesian Methods and Reliability Growth", AD-653460, March 1967.

41. Lieberman, Gerald J. "Some Problems in Reliability Estimation", Proceedings Third Annual Aerospace Reliability and Maintainability Conference, Washington, D.C. 1964, P.136-140.

42. MacFarland, W.J. "Uses of Bayes Theorem in Its Discrete Formulation for Reliability Estimation Purposes", Annals of Assurance Sciences, Vol 1 No. 2, 1968.

43. Martel, R.J. "Optimum Bayes Strategies in Truncated Life Testing", Presented at 27th National Mtg. Operations Res. Soc. of America, May 1965.

44. Mastran, D.V. "A Bayesian Approach for Assessing the Reliability of Air Force Re-Entry Systems", Annals of Assurance Sciences, Vol 1 No. 2, 1968.

45. Middleton, D. "A Structure for the Bayes Analysis of the Performance of Detection Systems with Multiple Sensors and Sites", Rand Corp., March 1967.

46. Morris, C. "Admissible Bayes Procedures and Classes of Epsilon Bayes Procedures for Testing Hypotheses in a Multinomial Distribution", AD 642 636, August 1966.

47. Olsson, J.E. "Implementation of a Bayesian Reliability Measurement Program", Annals of Assurance Sciences, Vol 1 No. 2, 1968.

48. Pierce, Don A. and Folks, J. Leroy, "A Method for Determining the Sensitivity of a Bayes Procedure to the Prior Distribution".

49. Pollock, S.M. "A Bayesian Reliability Growth Model", Postgraduate Thesis U.S. Naval Postgraduate School, May 1966.

50. Porter, Samual A. "Reliability Programs for Aerospace Systems and the Bayes Theorem to Assure Reliability", AD-664586, 10 August 1967.

51. Pozner, A.N. "A New Reliability Assessment Technique", American Society for Quality Control, 20th Annual Technical Conference, June 1966.

52. Press, S.J. "Statistical Estimation by the Empirical Bayes Method: Some Extensions and Logistical Applications", Written Under Contract AF 49 (638)-700; Proj. Rand, June 1965.

53. Pugh, E. "The Bayesian Approach to the Reliability - Confidence Relation for Exponential Failures", J. Operations Res. Soc., Vol 8 September/ October 1960.

54. Rutherford, J.R. and Krutchkoff, R.G., "The Emperical Bayes Approach: Estimating the Prior Distribution", Biometrika 54, 1967, P.326-328.

55. Savage, L.J. "Bayesian Statistics", Recent Developments in Information and Decision Processes, Ed. by Mochol, R.D. and Grey, P. 1961.

56. Schafer, R.E. "Bayesian Operating Characteristic Curves for Reliability and Quality Sampling Plans", Proceedings 10th Symposium on Reliability and Quality Control, 1964.

57. Schafer, R.E. "Bayes Single Sampling Plans by Attributes Based on Posterior Risks", Naval Research Logistics Quarterly, March 1967.

58. Schafer, R.E. "Reduction of Sample Sizes by Using Prior Quality and Reliability History", Proceedings of the Annual Meeting of the American Society for Quality Control, 1966.

59. Schafer, R.E. and Singpurwalla, N.D. "A Sequential Bayes Procedure for Reliability Demonstration", Annals of Assurance Sciences, Vol 8, 1969.

60. Schick, G.J. "Bayesian Concepts for Reliability and Confidence", Annals of Assurance Sciences, Vol 1, No. 1, 1968.

61. Schulhof, R.J. and Lindstrom, D.L. "Application of

Bayesian Statistics in Reliability", 12th National Symposium on R & QC, January 1966.

62. Shah, Haresh C. "Reliability Estimates from Run-Out or Non-Failure Data Using Information Theory", American Society for Quality Control Annual Technical Conference, 1968, P.425-434.

63. Soland, R.M. "Bayesian Analysis of the Weibull Process with Unknown Scale Parameter", Written Under Contract DA-44-188-ARO-1, August 1966.

64. Soland, R.M. "Bayesian Analysis of the Weibull Process with Unknown Scale Parameter and Its Application to Acceptance Sampling", IEEE Transactions on Reliability, June 1968.

65. Springer, M.D. and Thompson, W.E. "Bayesian Confidence Limits for the Reliability of Cascade Exponential Subsystems", IEEE Transactions on Reliability, September 1967.

66. Springer, M.D. and Thompson, W.E. "Bayesian Confidence Limits for System Reliability", Annals of Assurance Sciences, Vol 8, 1969.

67. Springer, M.D. and Thompson, W.E. "Bayesian Confidence Limits for Reliability of Redundant Systems when Tests are Terminated at First Failure", Technometrics, Vol 10, February 1968.

68. Typaldos, Z.A. and Brimley, D.E. "Point Estimation of Reliability from Results of a Small Number of Trials", Rand Memo RM-3044 (1962) Astia 276 150.

69. Von Mises, R. "On the Correct Use of Bayes' Formula", Ann. of Mathematical Statistics, Vol 13, 1942.

70. Wadsworth, H.M. and Gilbreath, S.G. "A Bayesian Procedure for the Design of Sequential Sampling Plans", Proceedings of ASQC, May 1967.

71. Wallenius, K.T. "Sequential Reliability in Finite Lots", Technometrics; Vol 11, No. 1, February 1969.

72. Weir, W.T. "Proceedings of the Second Missile and Space Division Seminar on Bayes Theorem and

Its Application to Reliability Measurement", AD-481 645, December 1965.

73. Weir, W. T. "Bayesian Reliability Evaluation", Proceedings of ASME R&M Symposium 1968, 384-386.

74. Weir, W.T. "On Bayes' Theorem and Its Application to Reliability Evaluation", Presented at 3rd Sem. on Bayes Theorem, April 27, 1967.

75. Weir, W.T. "Bayesian Reliability Evaluation", Annals of Assurance Sciences, Vol 1 No. 2, 1968.

76. Wetherill, G.B. "Bayesian Sequential Analysis", Biometrika, Vol 48, 1961.

77. Wetherill, G.B. "Some Remarks on the Bayesian Solution of the Single Sample Inspection Scheme", Technometrics, Vol 2 No. 3, 1960.

78. Yereance, R.A. "Reliability Facts and Factors -- Bayesian Statistics", Systems Design, Vol 9, April 1965.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Hughes Aircraft Company Systems Effectiveness Department Box 3310, Fullerton CA 92634	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

BAYESIAN RELIABILITY DEMONSTRATION: PHASE I - DATA FOR THE A PRIORI DISTRIBUTION

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Final Report August 1968 - December 1969

5. AUTHOR(S) (First name, middle initial, last name)
R. L. Schafer
John Collins
M. L. Luden, et al

6. REPORT DATE February 1970	7a. TOTAL NO. OF PAGES 140	7b. NO. OF REFS 78

8a. CONTRACT OR GRANT NO. F30602-69-C-0042	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. 5519	
c. Task No. 551902	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) RADC-TR-69-389
d.	

10. DISTRIBUTION STATEMENT
This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of RADC (EMNRR) GAFB, N. Y. 13440.

11. SUPPLEMENTARY NOTES PROJECT ENGINEER Anthony J. Feduccia(EMNRR) AC 315 330-4…	12. SPONSORING MILITARY ACTIVITY Rome Air Development Center (EMNRR) Griffiss Air Force Base, New York 13440

13. ABSTRACT

This final report is a result of a study performed for RADC under Contract Number F30602-69-C-0042. The purpose of the study was to fit one or more prior distributions to θ = MTBF = Mean Time Between Failure. In particular, the objectives were three: (a) establish criteria for data that would be suitable for fitting prior distributions to θ = MTBF; (b) develop methods of fitting and fit one or more prior distributions; and (c) perform robustness analysis of fitted prior distributions.

It was discovered that if the number of identical equipments and number of failures observed per equipment are relatively small, special methods of fitting are required. For the data used in this study, the inverted gamma distribution turned out to be a good prior fit.

Best Available Copy

DD FORM 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Reliability Testing Statistics Electronic Equipment						

UNCLASSIFIED
Security Classification

AFLC—Griffiss AFB NY 6 Mar 70-123